



Universidad Nacional Mayor de San Marcos

Universidad del Perú. Decana de América

Facultad de Ciencias Biológicas

Escuela Profesional de Genética y Biotecnología

**Caracterización molecular de 129 accesiones de quinua
(*Chenopodium quinoa* Willd.) de la región puno
mediante marcadores microsatélites**

TESIS

**Para optar el Título Profesional de Biólogo Genetista
Biotecnólogo**

AUTOR

Renzo Guillermo CÁRDENAS CÓRDOVA

ASESOR

Alberto Ernesto LÓPEZ SOTOMAYOR

Lima, Perú

2017



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

Referencia bibliográfica

Cárdenas, R. (2017). *Caracterización molecular de 129 accesiones de quinua (Chenopodium quinoa Willd.) de la región puno mediante marcadores microsatélites*. [Tesis de pregrado, Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Biológicas, Escuela Profesional de Genética y Biotecnología]. Repositorio institucional Cybertesis UNMSM.



UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS
(Universidad del Perú, DECANA DE AMÉRICA)

FACULTAD DE CIENCIAS BIOLÓGICAS ✓

ACTA DE SESIÓN PARA OPTAR AL TÍTULO PROFESIONAL DE
BIÓLOGO GENETISTA BIOTECNÓLOGO
(MODALIDAD: SUSTENTACIÓN DE TESIS)

Siendo las 10:12 horas del 11 de setiembre de 2017 en el Salón de Grados de la Facultad de Ciencias Biológicas y en presencia del jurado formado por los profesores que suscriben, se dio inicio a la sesión para optar al Título Profesional de Biólogo Genetista Biotecnólogo de RENZO GUILLERMO CÁRDENAS CÓRDOVA.

Luego de dar lectura y conformidad al expediente N° 007-EPGB-2017, el titulando expuso su tesis: "CARACTERIZACIÓN MOLECULAR DE 129 ACCESIONES DE QUINUA (*Chenopodium quinoa* WILLD.) DE LA REGIÓN PUNO MEDIANTE MARCADORES MICROSATÉLITES" y el Jurado efectuó las preguntas del caso calificando la exposición con la nota 19, calificativo: *Aprobado con excelentes honores*

Finalmente, el expediente será enviado a la Escuela Profesional de Genética y Biotecnología, y al Consejo de Facultad para que se apruebe otorgar el Título Profesional de Biólogo Genetista Biotecnólogo a RENZO GUILLERMO CÁRDENAS CÓRDOVA se eleve lo actuado al Rectorado para conferir el respectivo título, conforme a ley.

Siendo las 11:32 horas se levantó la sesión.

Ciudad Universitaria, 11 de setiembre de 2017.

Dr. ARMANDO YARLEQUE CHOCAS
(PRESIDENTE)

Blgo. ALBERTO LOPEZ SOTOMAYOR
(ASESOR)

Mg. MERY SUNI NINATAIPE
(MIEMBRO)

Blgo. RAFAEL LA ROSA LOLI
(MIEMBRO)

DEDICATORIA

*Dedicado a aquellas personas que aún no descubren
todo su potencial, a aquellos que valen un mundo pero
que aún no lo saben, ustedes pueden con las
adversidades del día al día, tómalos con calma, respira y
verás que nada es imposible, descubrirás que no hay
nada mejor que amarse a uno mismo.*

AGRADECIMIENTOS

Durante el desarrollo de esta tesis muchas personas me ayudaron a no darme por vencido con el fin de culminar este arduo trabajo.

Agradezco a mi madre Sandra Córdova quien durante todos estos años ha estado mejorando día a día, sanando de a pocos y probando que es una mujer fuerte, decidida y aguerrida, a mi padre Guillermo Cárdenas quien me ha enseñado que la vida no es fácil pero que estará ahí para apoyarme en lo que pueda y a mi hermano quien poco a poco me ha ido mostrando su amor incondicional. Les amo demasiado.

A mi tía Mary, a quien amo con todas mis fuerzas, tú creíste en mí desde que me regalaste mi primer libro de animales y seguiste creyendo en mí conforme pasaban los años, muchas gracias tía.

A Carlita, te quiero prima, gracias por esos consejos y por inspirarme a seguir y seguir con este trabajo, así como tú lo hiciste con el tuyo.

A mis asesores Alberto Lopez y Fernando Serna, ustedes han estado pendientes de que no haya desistido de esta investigación, gracias por eso.

A Claudia, Wendy, Roger, Paul, Andrea y Hiromi, compañeros del INIA, gracias por los consejos para mi investigación y por tenerme paciencia en el laboratorio. A la señora Clara e Iris, muchas gracias por su ayuda y orientación en el laboratorio.

A mis amigos de base los “muppets”: Dalia, Yaji, Henry y Eva. Ustedes estuvieron conmigo en mi etapa universitaria y post-universitaria demostrándome que estarán ahí para los momentos felices y tristes de mi vida, ahora sé que puedo contar con ustedes. Dalia, me has ayudado tanto en estos últimos 9 años en casi todo aspecto de mi vida, muchas gracias hermana. A Indira, Wendy y Lenin, gracias por su amistad, cariño y apoyo en estos últimos años.

A mi profesora y entrenadora Maritza Gutierrez, alguien a quien admiro mucho, quiero mucho y que me ha enseñado el poder que tiene uno mismo para controlar su vida, para amarse hasta más no poder y para enfrentar y afrontar todas las situaciones

que se presenten, muchísimas gracias miss Mary. A mis compañeros de la selección de aeróbicos quienes estuvieron conmigo enfrentando duros retos físicos demostrando lo que valemos y lo que somos capaces de hacer. Gracias Karinita, me escuchaste en esos momentos que necesitaba ser escuchado, sabes que puedes contar conmigo. Gracias Luisiño por ayudarme a no rendirme y preocuparte porque termine este trabajo.

A Piero, Jonas, Kike, Julian y Javier, gracias por esos momentos de diversión y alegría, ustedes hicieron este viaje más entretenido. Gracias Javier, me ayudaste a crecer como persona y ver que soy capaz de muchas cosas, así como tu también lo eres.

A mis amigos del colectivo Versiones, ustedes me enseñaron que ser diferente está bien y a valorarme por esas diferencias. Gracias Fabrizio por ayudarme en esos momentos de estrés y ansiedad, *slay the runway*.

Finalmente, expresar mi más sincero agradecimiento a todos lo que me apoyaron y creyeron en mí durante estos dos últimos años, ha sido un camino largo y difícil, pero de mucho aprendizaje y lecciones.

ABREVIATURAS

A	Diversidad Alélica
A _e	Número efectivo de alelos
ADN	Ácido desoxirribonucleico
AFLP	Amplified fragment length polymorphism (Polimorfismo en la longitud de los fragmentos amplificados)
AME	Approximate maximum emission (Emisión máxima aproximada)
AMOVA	Analysis of molecular variance (Análisis de varianza molecular)
BES-SSR	Bacterial artificial chromosome - end sequences – SSR (SSR basados en secuencias terminales de cromosomas artificiales bacterianos)
BIC	Bayesian information criterion (Criterio de información bayesiano)
BSA	Bovine serum albumin (Albúmina sérica bovina)
CCD	Charged-coupled device (Dispositivo con carga acoplada)
Cm	Centímetro
DAPC	Discriminant analysis of principal components (Análisis discriminante de componentes principales)
dNTPs	Deoxyribonucleotide triphosphates (Desoxirribonucleótidos trifosfato)
EDTA	Ethylendiaminetetraacetic acid (Ácido etilenodiaminotetracético)
EEA	Estación experimental agraria
EST-SSR	Expressed sequence tag - SSR (SSR basados en marcadores de secuencia expresada)
F	Coeficiente de endogamia
F _{RT}	Proporción de la diferenciación total debido a diferencias entre las regiones

F_{SR}	Proporción de la diferenciación total debido a diferencias entre las subpoblaciones dentro de las regiones
F_{ST}	Proporción de la endogamia total de la población debido a la diferenciación entre subpoblaciones (Índice de fijación)
H	Diversidad génica (Nei, 1987)
H_e	Heterocigosidad esperada
H_{exp}	Heterocigosidad esperada por locus (adegenet)
H_l	Heterocigosidad observada subpoblacional promedio
H_o	Heterocigosidad observada
H_s	Heterocigosidad esperada (adegenet)
H_S	Heterocigosidad esperada subpoblacional promedio
H_T	Heterocigosidad total
IAM	Infinite alleles model (Modelo de los alelos infinitos)
INIA	Instituto Nacional de Innovación Agraria
LOD	Limit of detection (Límite de detección)
m	Metros
MCMC	Markov chain Monte Carlo (Cadenas de Markov Monte Carlo)
ml	Mililitros
mm	Milímetros
mM	Milimolar
msnm	Metros sobre el nivel del mar
ng	Nanogramos
nm	Nanómetros
NMDS	Non-metric multidimensional scaling (Escalado multidimensional no métrico)
pb	Pares de base
PCA	Principal Component Analysis (Análisis de componentes principales)

PCoA	Principal Coordinates Analysis (Análisis de coordenadas principales)
PCR	Polymerase chain reaction (Reacción en cadena de la polimerasa)
PHR	Peak height ratio (relación de altura del pico)
PIC	Polymorphic information content (Contenido de información polimórfica)
POP	Performance optimized polymer (Polímero de acción optimizada)
QTL	Quantitative trait locus (Locus de carácter cuantitativo)
RAPD	Random amplified polymorphic DNA (ADN polimórfico amplificado aleatoriamente)
RFU	Relative fluorescent units (Unidades de fluorescencia relativa)
R _{ST}	Proporción de variación en la longitud de los alelos que es debido a las diferencias entre las poblaciones
SD	Standard deviation (Desviación estándar)
SMM	Stepwise mutation model (Modelo mutacional por pasos)
SSR	Simple sequence repeats (Secuencias simples repetidas)
STR	Short tandem repeats (Secuencias cortas repetidas)
UHe	Unbiased heterozygosity (Heterocigosidad no sesgada)
ul	Microlitro
um	Micrómetro
uM	Micromolar
UPGMA	Unweighted pair-group method with arithmetic average (Método de agrupamiento por medias aritméticas no ponderables)
V	Voltios

ÍNDICE

DEDICATORIA.....	I
AGRADECIMIENTOS	II
ABREVIATURAS	IV
ÍNDICE.....	VII
RESUMEN	IX
ABSTRACT.....	X
1. INTRODUCCIÓN	1
2. MARCO TEÓRICO.....	3
2.1. La especie <i>Chenopodium quinoa</i> Willd.	3
2.1.1. Origen, historia y domesticación	3
2.1.2. Posición Taxonómica	6
2.1.3. Características botánicas y morfológicas	6
2.1.4. Distribución Geográfica y tipología actual	9
2.1.5. Citogenética	13
2.1.6. Valor nutricional y económico	13
2.2. Marcadores moleculares microsatélites	16
2.2.1. Composición y distribución de los microsatélites.....	16
2.2.2. Clasificación de los microsatélites	17
2.2.3. Ventajas y limitaciones de los marcadores microsatélites	18
2.2.4. Aplicaciones de los microsatélites	22
2.2.5. Trabajos previos en quinua utilizando microsatélites.....	23
2.3. Métodos de separación, detección y marcaje del ADN	31
2.3.1. Electroforesis capilar	31
2.3.2. Detección y marcaje del ADN por fluorescencia.....	32
2.4. La diversidad genética y su análisis	34
2.4.1. Diversidad Genética.....	34
2.4.2. Análisis de la diversidad genética	36
3. HIPÓTESIS.....	44
4. OBJETIVOS.....	44
4.1. Objetivo general	44
4.2. Objetivos específicos	44
5. MATERIALES Y MÉTODOS	45
5.1. Material vegetal.....	45
5.2. Extracción del ADN genómico.....	45
5.3. Cuantificación del ADN genómico	47

5.4.	Selección de marcadores microsatélites y amplificación por PCR.....	48
5.5.	Separación de productos amplificados por electroforesis capilar	49
5.6.	Genotipado de los productos amplificados	50
5.7.	Análisis de datos	51
5.7.1.	Determinación de grupos genéticos	51
5.7.2.	Asignación de alelos a isoloci y determinación del tipo de herencia	53
5.7.3.	Estimación de frecuencias alélicas e índices de diversidad genética	54
5.7.4.	Distancia genética, índices de diferenciación genética y análisis de coordenadas principales	55
6.	RESULTADOS	56
6.1.	Calidad del ADN genómico	56
6.2.	Amplificación de las regiones microsatélites y separación de los productos amplificados.....	56
6.3.	Estructura poblacional.....	56
6.4.	Determinación del tipo de herencia	58
6.5.	Índices de diversidad y diferenciación genética.....	59
6.6.	Distancia de Bruvo y análisis de coordenadas principales	60
7.	DISCUSIÓN	61
7.1.	Calidad de la extracción del ADN genómico.....	61
7.2.	Dosis alélica y frecuencias alélicas	61
7.3.	Disómica vs. Polisómica.....	64
7.4.	Estructura poblacional.....	66
7.5.	Diversidad Genética	68
8.	CONCLUSIONES	73
9.	RECOMENDACIONES	74
10.	REFERENCIAS BIBLIOGRÁFICAS.....	75
11.	GLOSARIO.....	92
12.	ANEXOS	96
12.1.	Tablas.....	96
12.2.	Figuras	117

RESUMEN

La quinua (*Chenopodium quinoa* Willd.) es un cultivo importante debido a las características nutricionales favorables de su semilla. Actualmente en el Perú existen 8 bancos de germoplasma con 6302 accesiones de quinua. El estudio de la diversidad genética de la quinua permite conocer su potencial para adaptarse a las condiciones adversas del ambiente además de ayudar a identificar variedades con características agronómicas óptimas. El objetivo de esta tesis fue caracterizar a nivel molecular la diversidad genética de 129 accesiones de quinua de la región Puno. Se evaluó la diversidad genética y estructuración poblacional utilizando 15 marcadores microsatélites mediante un sistema de PCR multiplex y electroforesis capilar. Se detectaron un total de 179 alelos entre las accesiones, en un rango de 7 a 19 alelos por locus con un promedio de 11.93, mientras que la heterocigosidad esperada total fue de 0.77. Se detectaron 3 grupos genéticos (K1, K2, K3) cuyos niveles de diferenciación genética entre los grupos fue muy baja ($F_{ST} = 0.028 - 0.046$) por lo que no se formaron grupos discretos según los análisis del PCoA y del DAPC. No se encontraron accesiones duplicadas en las 129 accesiones. El análisis bioinformático sugirió que la quinua presenta una herencia polisómica por lo que fue tratada como un organismo autopoliploide. Los resultados demostraron que las 129 accesiones presentaron una alta diversidad genética sin una estructura poblacional. Este trabajo constituye uno de los primeros pasos para la aplicación de métodos orientados al buen manejo del banco de germoplasma o al fitomejoramiento.

Palabras clave: microsatélite, Chenopodium quinoa, diversidad genética, estructura poblacional, electroforesis capilar.

ABSTRACT

Quinoa (*Chenopodium quinoa* Willd.) is an important seed crop due to the beneficial nutritional traits. In Peru, there are 8 genebanks conserving 6302 quinoa accessions. The study of the genetic diversity of quinoa allows to know its potential to adapt to the adverse conditions of the environment besides helping to identify varieties with optimal agronomic traits. The objective of this thesis was to characterize to a molecular level 129 quinoa accessions from Puno. Genetic diversity and population structure were evaluated using 15 microsatellite markers through multiplex PCR and capillary electrophoresis. A total of 179 alleles were detected among the quinoa accessions, ranging from 7 to 19 alleles per locus with an average of 11.93, whereas the total expected heterozygosity was 0.77. Three genetic groups were detected (K1, K2, K3) whose genetic differentiation levels between groups were very low ($F_{ST} = 0.028 - 0.046$), so there were not discrete groups according to the PCoA and DAPC analysis. No duplicated accession was found within the 129 accessions. Bioinformatic analysis suggested polysomic inheritance in quinoa therefore it was treated as an autopolyploid organism. The results showed that the 129 accessions presented a high genetic diversity without any population structure. This work is one of the first steps for the applications of methods oriented to the good genebank management or plant breeding.

Keywords: microsatellite, Chenopodium quinoa, genetic diversity, population structure, capillary electrophoresis.

1. INTRODUCCIÓN

La quinua (*Chenopodium quinoa* Willd.) es una planta herbácea y dicotiledónea cuya semilla es importante debido a sus características nutricionales ya que posee los aminoácidos esenciales para el ser humano (Bhargava y Srivastava, 2013). Esta planta tiene una distribución geográfica en Sudamérica a lo largo de los Andes, extendiéndose desde el sur de Colombia a 5° latitud norte hasta la décima región de Chile a 43° latitud sur, a alturas que van desde el nivel del mar como por ejemplo las tierras bajas de Chile hasta los 4000 msnm en zonas del altiplano de Perú y Bolivia (Bhargava y Srivastava, 2013). Su domesticación ocurrió en los Andes centro-sur antes de los 3000 A.C. y su cultivo se remonta desde el periodo Formativo hasta el Periodo Inca, donde ya era base de la subsistencia de las sociedades (Planella et al., 2013). La quinua estuvo sujeta a un intenso proceso de mejoramiento debido a la enorme variación fenotípica que se puede encontrar alrededor del lago Titicaca, zona donde también se encuentra una alta diversidad genética (Christensen et al., 2007).

La diversidad genética es importante porque permite a las poblaciones adaptarse a los cambios ambientales y es una fuente importante de biodiversidad. La diversidad genética puede ser cuantificada utilizando herramientas biotecnológicas como los marcadores moleculares que son fragmentos de ADN polimórficos entre diferentes individuos (Jiang, 2013). Entre los distintos tipos de marcadores moleculares, los microsatélites o SSR se han convertido en una herramienta poderosa para el análisis de diversidad dado que son altamente polimórficos, multialélicos, codominantes, altamente reproducibles, distribuidos ampliamente y de manera uniforme en el genoma (Fuentes et al., 2012).

Actualmente en el Perú existen 8 bancos de germoplasma con 6302 accesiones de quinua. Este tipo de conservación *ex situ* complementa a la conservación *in situ* que se da en los mismos campos realizado por los campesinos y tiene como objetivo salvaguardar la diversidad genética poniéndola a disposición de los fitomejoradores (Rojas et al., 2014). La conservación *ex situ* de la quinua puede ayudar a garantizar la

seguridad alimentaria del país, por lo que la correcta caracterización morfológica, agronómica, bioquímica y molecular de las accesiones es de carácter prioritario para estos bancos. La caracterización molecular de las accesiones de los bancos de germoplasma constituye uno de los primeros pasos para la aplicación de métodos orientados al buen manejo de los bancos, como la formación de colecciones núcleo, o al fitomejoramiento mediante la selección asistida por marcadores (MAS) (Kalia et al., 2011).

Actualmente se tienen 129 accesiones de quinua provenientes de la región Puno y conservadas en el banco de germoplasma del INIA en Puno (Ilpa) que cuentan con la información agronómica respectiva pero no se tiene conocimiento del nivel de diversidad genética ni la estructura poblacional.

Por tal motivo, el presente trabajo se abocó a la caracterización molecular de estas 129 accesiones de quinua del altiplano que presentaban mayor información de caracterización agronómica, utilizando marcadores microsatélites que ya han sido aplicados en una parte de la colección nacional (Via y Rada, 2015).

2. MARCO TEÓRICO

2.1. La especie *Chenopodium quinoa* Willd.

2.1.1. Origen, historia y domesticación

El origen alotetraploide de la quinua ya ha sido sugerido en diferentes estudios mediante análisis genético, molecular y citogenético (Kolano et al., 2016; Kolano et al., 2011; Maughan et al., 2006; Ward, 2000b; Wilson, 1990). Los progenitores de *C. quinoa* son dos diploides ancestrales del género *Chenopodium sensu stricto* que vienen de dos diferentes linajes (Figura 2.1): el primer linaje incluye diploides americanos (por ejemplo *C. neomexicanum*, *C. incanum*) los cuales poseen el genoma A, y el segundo linaje incluye diploides del viejo mundo (por ejemplo *C. ficifolium*) que poseen el genoma B (Kolano et al., 2015). Las especies parentales que dieron origen a la quinua aún no han podido ser identificadas (Kolano et al., 2016).

En cuanto a su historia, la quinua como grano alimenticio ha sido considerado por siglos como un cultivo importante en la región de los Andes, no es sorpresa que la palabra quinua en quechua signifique “grano madre”, y la quinua ha tenido tantos nombres como regiones que la conocían. Los chibchas en Colombia la denominaban “pasca” que significa “la olla o comida del padre”; en el idioma aymara la quinua ha tenido varios nombres según la variedad: la morada se llamaba “cami”, la blanca y la más apreciada “ppfique”, la colorada “kana llapi”, la amarilla “cchusllunca” y la silvestre “isualla”; en el norte de Chile se cultivaba quinua y se le denominada en atacameño “dahue” (Tapia, 1979).

Los antiguos agricultores de los Andes fueron los que empezaron la domesticación de la quinua y la transición de la forma silvestre a la doméstica en un periodo de 5000 años, agricultores de diferentes civilizaciones antiguas en diferentes tiempos y diferentes zonas geográficas incluyendo Perú (5000 a.C.), Chile (3000 a.C.) y Bolivia (750 a.C) (Kadereit et al., 2003). En el norte del Perú, el cultivo de quinua fue común, pero en asociación con el maíz. Más al sur, está alcanzó importancia tanto en el Callejón de

Huaylas como en el valle del Mantaro donde fue ampliamente cultivada por la tribu de los huancas (Tapia, 1979).

Durante este proceso, la quinua fue sujeta a varios procesos de selección por diferentes grupos humanos de diferentes culturas dentro de Sudamérica (Figura 2.2) como los chibchas, andakíes e inganos en el sur de Colombia; los aymaras y quechuas en Perú, Bolivia y norte de Chile; los diaguitas y calchaquíes en el norte de Argentina; y mapuches en el sur de Chile (Mujica, 2004). Estos procesos de selección permitieron la pérdida de características desfavorables para el cultivo de quinua como la presencia de un episperma grueso y alto contenido de saponina, y la ganancia de características favorables como semillas más grandes y resistencia al estrés abiótico (Figura 2.3). A pesar de que este proceso haya reducido la diversidad genética de la quinua en un sentido amplio, la diversidad fenotípica sigue siendo amplia y puede ser observada en la diversidad de colores de la planta y semilla, de formas de ramificación y de la panoja, de productividad, de resistencia a condiciones ambientales desfavorables y a enfermedades (Fuentes y Bhargava, 2011).

La historia natural y política sugiere que la diversidad genética de la quinua podría haber pasado al menos por tres eventos genéticos de cuello de botella, los cuales no han representado una gran pérdida para la diversidad genética.

El primero y potencialmente el más severo de los cuellos de botella pudo haber ocurrido cuando los dos ancestros diploides de la quinua se hibridaron, asumiendo que éste resultó en (1) una especie alotetraploide aislada reproductivamente, probablemente *C. berlandieri* (Wilson, 1980), y (2) que este evento ocurrió solo una vez, lo que significó que el complejo de especies tetraploides del nuevo mundo tuvieron un origen filogenético.

El segundo cuello de botella putativo pudo haber ocurrido cuando la quinua fue domesticada a partir de sus ancestros tetraploides silvestres, aunque este cuello de botella no haya representado un evento muy fuerte debido a la capacidad de la quinua actual para hibridarse con quinuas silvestres u otras especies silvestres del género

Chenopodium, como *C. quinoa* var. *melanospermum* o *C. hircinum* respectivamente. Además, existe otra posibilidad planteada sobre la base de estudios de diversidad utilizando marcadores de ADN en donde se sugiere que la quinua ha sido domesticada 2 veces: una en el altiplano, y la segunda en la costa de Chile (Christensen et al., 2007; Fuentes et al., 2009a).

El tercer cuello de botella sucedió hace más de 400 años, desde el periodo de la conquista hasta la década de los 80 debido a la marginalización del cultivo de quinua por razones culturales (Fuentes et al., 2009b; Jellen et al., 2011).

En cuanto al centro de diversidad genética de la quinua, dos estudios publicados en 1979 y 1988 (Gandarillas, 1979b; Wilson, 1988) identificaron al altiplano boliviano como el centro de diversidad genética de la quinua. Posteriormente Christensen y colaboradores usando marcadores SSR identificaron el centro de diversidad genética en el área del altiplano entre Perú y Bolivia mientras que en las regiones de Ecuador y Argentina presentaban limitada diversidad (Christensen et al., 2007). Los recientes análisis genéticos son consistentes con la idea de que la quinua ha existido hasta ahora como dos distintos acervos genéticos: la quinua del altiplano asociados con sus parientes silvestres (quinua “ajara” o “ashpa”, *C. quinoa* ssp. *milleanum* Aellen, también referido como *C. quinoa* var. *melanospermum* Hunziker), y la “kinwa” de la cultura mapuche en el centro y sur de la costa chilena representando a un segundo centro de diversidad de la quinua (Jellen et al., 2011). La presencia de quinua en tierras más sureñas ha sido explicada por el intercambio de semillas entre los incas y grupos aborígenes que vivían en zonas al sur del altiplano como el altiplano chileno (17 °S), isla Chiloé (42 °S) y Puerto Río Tranquilo (47 °S) (Bazile et al., 2013).

Además, la especie *Chenopodium hircinum*, pariente silvestre de la quinua de las tierras bajas de Argentina, puede ser considerada como un tercer acervo genético representando un remanente del cultivo ancestral de quinua en esa parte de Sudamérica (Wilson, 1990).

2.1.2. Posición Taxonómica

Según el Sistema de Información Taxonómico Integrado (ITIS, 2011), la jerarquía taxonómica de *Chenopodium quinoa* Willd. es la siguiente:

Reino	Plantae
Subreino	Viridiplantae
Infrarreino	Streptophyta
Superdivisión	Embryophyta
División	Tracheophyta
Subdivisión	Spermatophytina
Clase	Magnoliopsida
Superorden	Caryophyllanae
Orden	Caryophyllales
Familia	Amaranthaceae
Género	<i>Chenopodium</i> L.
Especie	<i>Chenopodium quinoa</i> Willd.

Al nivel de familia, Amaranthaceae y Chenopodiaceae han sido tratadas durante mucho tiempo como familias cercanamente relacionadas. Estudios basados en morfología, anatomía y fitoquímica de las dos familias revelaron un número de características en común. Estos estudios en conjunto con los datos moleculares (Rettig et al., 1992; Downie et al., 1997; Singh, 2010) apoyan la combinación de las dos familias en Amaranthaceae. El status independiente de Chenopodiaceae y Amaranthaceae dejó de existir en 1998 cuando el APG (Angiosperm Phylogeny Group) publicó el APG I (Angiosperm Phylogeny Group, 1998) en donde se combinaron las dos familias en una sola (Amaranthaceae) sobre la base de las evidencias morfológicas y moleculares.

2.1.3. Características botánicas y morfológicas

La quinua presenta una raíz pivotante que a partir de unos pocos centímetros del suelo puede ramificarse en raíces secundarias y terciarias, y dependiendo de la altura de la planta la raíz puede llegar a profundidades cerca de la superficie (12.6 – 15 cm) o profundidades de 1.5 m (Bhargava y Srivastava, 2013). Su sistema radicular muy

ramificado hace que la especie sea más resistente a la sequía y fuerte contra los vientos (Gandarillas, 1979a).

El tallo es cilíndrico cerca de la superficie del suelo y se vuelve anguloso debido a la presencia de hojas alternas en sus cuatro caras. La altura puede variar entre 50 cm y 2 m dependiendo de la variedad. La textura de la médula es blanda en plantas jóvenes y se vuelve esponjosa, hueca y de color crema cuando va madurando; por otro lado, la corteza es firme y compacta, formada por tejidos fuertes (Gandarillas, 1979a). Puede ser de color verde, púrpura, amarillo o rojo (Lescano, 1981). El color rojo es debido a la presencia de betacianinas, un tipo de betalaínas (Gallardo et al., 2000). El hábito del tallo (Figura 2.4) puede ser sencillo, cuando las ramas son poco desarrolladas y alcanzan unos pocos centímetros; o ramificado, donde las ramas son tan largas que pueden llegar a la altura de la panoja principal formando otras panojas o bien crecen por los lados dando a la planta el aspecto de árbol de navidad (Gandarillas, 1979a).

Las hojas de *C. quinoa*, como toda dicotiledónea, presentan peciolo y láminas. El peciolo es largo, fino, acanalado en su lado superior y su largo varía dependiendo de donde se originen. Los que se originan directamente del tallo son más largos que los que se originan de las ramas primarias y secundarias. La lámina es polimórfica en la misma planta, siendo lanceoladas o triangulares las hojas superiores y romboidales o triangulares las hojas inferiores. La lámina por lo general es plana, aunque existen individuos de láminas onduladas. Las hojas jóvenes pueden presentar papilas globosas de 1.4 mm de diámetro que pueden ser blancas, púrpuras o rojas, y encontrarse en ambas caras de la hoja. Estas papilas también se pueden encontrar en tallos jóvenes y en las inflorescencias. El borde de las hojas puede ser liso, dentado o aserrado. El número de dientes de la hoja varía según la raza, de 3 a 20 dientes, y también varía según la localización. Las razas del centro-norte del Perú y Ecuador tienen hojas más aserradas, mientras que las razas bolivianas tienen pocos dientes o en algunos casos ninguno (Gandarillas, 1979a). El color de las hojas de plantas jóvenes es generalmente verde, pero en plantas maduras hay hojas amarillas, púrpuras o rojas; incluso puede

haber coloración roja en las venas y en la base de los peciolos (Bhargava y Srivastava, 2013).

La inflorescencia es una panoja muy ramificada de 15 a 70 cm de longitud y surge de la parte superior de la planta (inflorescencia terminal) y de las axilas de las hojas inferiores (inflorescencias axilares). Tiene un eje principal anguloso y con dos surcos paralelos en cada cara, por donde nacen las ramas secundarias y terciarias. Las inflorescencias axilares presentan un crecimiento determinado, las cuales dejan de crecer una vez que aparecen las flores hermafroditas. Hay ramas cortas de 0.5 a 3 cm en donde se agrupan las flores en número de 20 a más sobre un receptáculo, estas ramas con flores se denominan glomérulos. Las ramas secundarias y terciarias también presentan una flor hermafrodita terminal. Los glomérulos que se agrupan a lo largo de los ejes principal o secundarios dan lugar a las dos formas principales de inflorescencias: la glomerulada es aquella donde los glomérulos están insertados en las ramas terciarias, y la amarantiforme donde los glomérulos están insertados directamente en ramas secundarias (Figura 2.5). La inflorescencia ancestral es la glomerulada, que es de carácter dominante sobre la inflorescencia de tipo amarantiforme, de carácter recesivo (Gandarillas, 1979a).

Las flores de *C. quinoa* son incompletas, ya que carecen de pétalos. Las flores en el glomérulo pueden ser hermafroditas o pistiladas. Las flores hermafroditas, además de ser apicales, sobresalen de las pistiladas que se encuentran en la parte inferior del glomérulo. La flor hermafrodita está constituida por cinco sépalos, el gineceo con un ovario elipsoidal con dos o tres ramificaciones estigmáticas, y el androceo que rodea al gineceo y está formado por 5 estambres curvos y cortos. La flor femenina solamente consta de sépalos y gineceo. El tamaño del primero varía de 2 a 5 mm y del segundo de 1 a 3 mm. Los sépalos en ambas flores están cubiertos de papilas en el lado externo. Las flores pueden ser sésiles o pediceladas, pudiendo en algunos casos tener los pedicelos más de 5 mm (Gandarillas, 1979a).

El fruto es un aquenio cubierto por el perigonio, del que se desprende con facilidad al frotarlo cuando está seco. Su tamaño varía de 1.8 a 2.6 mm en diámetro y su forma (Figura 2.6) puede ser cilíndrica, cónica o elipsoidal (Bhargava y Srivastava, 2013). El color del fruto está dado por el perigonio y se asocia directamente con el de la planta pudiendo ser de color verde, púrpura o rojo. El pericarpio del fruto que está pegado a la semilla presenta alveolos y en algunas variedades se pueden separar fácilmente. El pericarpio contiene la saponina que le transfiere el sabor amargo a la semilla. El nivel de saponina varía en las diferentes variedades de quinua. Las semillas pueden tener forma cilíndrica, cónica o elipsoidal y pueden germinar muy rápidamente en unas pocas horas de haber sido expuestas a humedad. El peso de las semillas varía de 2 a 6 mg. La semilla está envuelta por el episperma en forma de una membrana delgada. El embrión está formado por los cotiledones y la radícula, y constituye la mayor parte de la semilla. El embrión envuelve al perisperma como un anillo blanco. El perisperma, compuesto de tejido de reserva, es almidonoso y normalmente de color blanco (Gandarillas, 1979a).

2.1.4. Distribución Geográfica y tipología actual

La quinua se distribuye a lo largo de la región andina, desde Colombia (Pasto) hasta el norte de Argentina (Jujuy y Salta) y Chile (Antofagasta), en donde se han observado cultivos de quinua creciendo en tierras a nivel de mar. De acuerdo a Rojas (Rojas, 1998), la distribución geográfica de la quinua se extiende desde los 5° N en el sur de Colombia hasta los 43° S en la décima región de Chile (Región de Los Lagos). La distribución natural es desde los 2° N hasta los 40° S (Fuentes y Bhargava, 2011). Las principales áreas del cultivo actual de quinua se extienden desde el extremo sur de Colombia a través de Ecuador, Perú y Bolivia, con extensiones en el altiplano chileno (este de Tarapacá) y el norte de Argentina (Jujuy y Salta) (Wilson, 1990). Su distribución altitudinal varía desde el nivel del mar en Chile hasta los 4000 m en el altiplano en Perú y Bolivia. En el Perú, algunas áreas de cultivo de quinua son Cajamarca, Callejón de

Huaylas en Ancash, el valle del Mantaro y tierras altas de Jauja en Junín, Andahuaylas en Ayacucho y en las tierras altas del departamento de Cuzco, pero es en el altiplano del Collao donde la quinua adquiere realmente importancia porque se concentra el mayor porcentaje de quinua cultivada a nivel nacional (Tapia, 1979). Puno es el principal departamento productor de quinua (75% de la producción total), seguido por Huancayo (10%) y Cuzco (5%) (Mujica et al., 2003).

La quinua presenta una extensa distribución porque es capaz de tolerar diferentes condiciones de humedad relativa (desde 40% hasta 88%), una amplia gama de altitudes (desde el nivel del mar hasta los 4000 m) y una variedad de temperaturas (de -8 °C a 38° C). Su increíble tolerancia se refleja en las diferencias que presenta a nivel fenológico, morfológico, comportamental, agrológico, ecológico y de resistencia a factores bióticos y abióticos. La diversidad de la quinua, a nivel continental, ha sido asociada con 5 principales ecotipos (Figura 2.7): el altiplano (Perú y Bolivia), valles interandinos (Colombia, Ecuador y Perú), salares (lagos salados; Bolivia, Chile y Argentina), yungas (Bolivia) y la costa (Chile); cada una de estas están asociadas con subcentros de diversidad que se originaron alrededor del Lago Titicaca (Risi y Galwey, 1984). Actualmente existen 8 tipos de quinua para describir su diversificación y utilización de acuerdo a sus zonas agroecológicas (Bazile et al., 2013):

- Quinua del altiplano (altiplano del norte): Esta zona, situada cerca del lago Titicaca y el altiplano del norte, representa una región relativamente templada y húmeda, siendo su precipitación anual promedio de unos 600 mm. Las quinuas son pequeñas, de diferentes colores, con un contenido variable de saponina y un periodo vegetativo de 6 meses. Producen semillas pequeñas a medianas, son menos resistentes a las heladas y sequías y pueden crecer ocasionalmente en suelos salinos. Tienen pocas ramas y una sola panoja con abundantes hojas. Son moderadamente resistentes al mildiu (*Peronospora farinosa* Fr.) y son atacados por orugas cortadoras de plantas jóvenes (*Feltia experta* Walker, *Eurisacca quinoa* Povolny) y por pájaros. Requieren una precipitación anual

entre 700 y 800 mm y crecen a una altitud de 3850 m. Algunas de las variedades de quinuas de este tipo son Kancolla, Blanca de July, Chullpi y Pasankalla en Puno (Perú) y La Paz (Bolivia).

- Quinua de los salares (altiplano del sur): Con condiciones de un típico desierto frío, esta zona se eleva a 3600 m y está rodeado de volcanes, tiene al salar de Uyuni en su centro. Las precipitaciones son de unos 350 mm en la parte norte del salar, y raramente excede los 150 mm en la parte sur, con más de 200 días de heladas al año. Estas plantas son grandes, ramificadas, de diferentes colores, con grandes granos de 2.2-2.9 mm con alto contenido de saponina, resistentes a las sequías, adaptados a suelos salinos y arenosos de las orillas de los salares. Como ejemplos, se tiene a las variedades Pandela, Utusaya, Toledo y Achachino.
- Quinuas de los valles interandinos: Estas zonas se encuentran a altitudes desde 2500 a 3200 m, con precipitaciones anuales que varían desde 800 hasta 900 mm, con suelos más fértiles y clima más cálido, aunque estas condiciones propicien los ataques por parásitos. Las plantas son altas y ramificadas, con hojas grandes, produciendo granos grandes y pequeños de diferentes colores, tienen un periodo vegetativo largo, son susceptibles al mildiu y pueden tener contenido alto o bajo de saponina. Algunas variedades representantes de quinuas de los valles interandinos son: Amarilla de Marangani, Blanca de Junín, Acostambo, Roja Coporaque, Nariño, etc.
- Quinuas de zonas áridas y condiciones secas (altiplano occidental): Estas zonas se caracterizan por estar a 3900 m con precipitaciones anuales de 150-350 mm y se localizan en partes de Perú, Bolivia y Chile. Las plantas son pequeñas con un periodo vegetativo corto debido a los 2 meses de lluvia anuales, con adaptaciones para soportar el estrés por sequía, con hojas pequeñas, colores variados, alto contenido de betacianina y oxalato de calcio, estructura profunda y altamente ramificada de las raíces, producen granos pequeños a medianos y

con un alto contenido de saponina. Algunos ejemplos de variedades son Antahuara, Ucha, Ccoyto y Roja Ayauchana.

- Quinuas de altitudes altas y clima fresco: Zonas de 4000 m de altitud con precipitaciones anuales de 800 mm y días cortos y soleados. Las plantas son pequeñas con colores vivos como amarillo, rojo o púrpura en las plantas mismas o en los granos, con panojas glomeruladas compactas, resistencia al frío y fuertes ventiscas y resistentes a la radiación ultravioleta. Los granos son amargos y con alto contenido proteico. Se tienen a variedades como Huariponcho, Witulla, Kellu, Kancolla y Roja.
- Quinuas de las regiones costeras y cerca del mar: Se encuentran en áreas con precipitaciones anuales de 500-600 mm y con suelos salados y arenosos. Son plantas medianamente ramificadas, con panojas glomeruladas, tolerantes a la sal con hojas pequeñas, con granos pequeños y duros y resistentes a la humedad excesiva. Algunas variedades de estas zonas son la quinua blanca, Kinwa mapuche, Lito, Faro e Islunga.
- Quinuas de la selva y zonas tropicales: Las zonas donde crecen estas quinuas tienen una altitud que varía entre los 800 y 1800 m, y presentan una precipitación anual de 1500 mm. Las plantas son altas, altamente ramificadas, con periodo vegetativo largo, hojas grandes, con colores intensos y brillantes, panojas grandes y sueltas (amarantiformes) y con granos pequeños. Son resistentes al mildiu, humedad excesiva, al calor y pueden crecer en suelos inundados.
- Quinuas de altas precipitaciones y zonas de humedad: Estas zonas presentan fuertes precipitaciones (2000-3000 mm). Son quinuas altas, altamente ramificadas, panojas grandes, con granos pequeños, de alto rendimiento y con un periodo vegetativo largo. Presentan un sistema radicular ancho y grueso para establecerse en suelos mal drenados. Son resistentes al mildiu, pero son fuertemente atacados por caracoles y babosas.

2.1.5. Citogenética

Estudios citológicos han establecido que *C. quinoa* tiene un número cromosómico de $2n = 4x = 36$ (Palomino et al., 1990; Bhargava et al., 2006a). Es un organismo aloploiploide, sus 36 cromosomas se pueden agrupar en 9 grupos (Catacora, 1977). Algunos hallazgos que apoyan la alotetraploidía son la duplicación del loci Lap (Wilson, 1976), herencia disómica de algunos caracteres (Simmonds, 1971), y las proporciones de segregación en la F1 y F2 las cuales indican herencia disómica-digénica y tetrasómica en algunos rasgos (Ward, 2000a). El cariotipo de la quinua presenta un par de cromosomas con satélites, estos cromosomas pueden ser medianos o mediano-submedianos (Bhargava et al., 2006a). La presencia de un par cromosómico con regiones satélites también ha sido corroborada por estudios de hibridación por fluorescencia *in situ* utilizando el ADNr 45S mostrando dos sitios de hibridación en dos cromosomas homólogos (Kolano et al., 2001). El cromosoma 1 puede ser mediano o mediano-submediano, los cromosomas más conservados (todos medianos) son los cromosomas 4, 9 y 18, y los cromosomas más variables son los cromosomas 10 y 13 (Bhargava et al., 2006a).

2.1.6. Valor nutricional y económico

El perisperma, embrión y endosperma son las tres áreas donde las reservas alimenticias se almacenan dentro del grano. El almidón es almacenado en el perisperma y los lípidos y proteínas en el endosperma y embrión (Bhargava et al., 2006b).

El contenido de proteína reportado en la literatura para el grano de quinua es de 12-23% (Abugoch, 2009). Comparando este grano andino con los cereales se puede observar que el contenido total proteico del grano de quinua (16.3% peso seco) es más alto que el de cebada (11%), arroz (7.5%) y maíz (13.4%). Su contenido se acerca bastante al del trigo (15.4%), aunque es menor que el de las semillas de leguminosas. Su balance de aminoácidos esenciales es excelente debido a su rango más amplio de aminoácidos en comparación a los cereales y legumbres, con contenidos altos de lisina (5.1-6.4%) y

metionina (0.4-1%). La calidad de las proteínas de quinua es tan alta como la de la caseína (Ranhotra et al., 1993), sin embargo, la digestibilidad de las proteínas disminuye cuando no se lava los granos debido a las saponinas (Ruales y Nair, 1992). Específicamente, las proteínas de almacenamiento de los granos son la fuente de proteínas consumidas directamente por los humanos, siendo las albúminas y globulinas las principales proteínas de almacenamiento (Fairbanks et al., 1990).

Con respecto al contenido de carbohidratos, el de la quinua es comparable con el de cebada y el de arroz. La cantidad de fibra dietética total de quinua es cercana al de los cereales (7-9.7% peso seco) y el contenido de fibra soluble está entre un 1.3% y 6.1% en peso seco. El almidón es el mayor componente dentro de los carbohidratos, está presente en un porcentaje entre 32% y 69.2% (Abugoch, 2009) y consiste de gránulos pequeños uniformes con diámetros menores que 3 μm (Atwell et al., 1983). Además, existe un 3% de azúcares simples y son principalmente maltosa, seguido de D-galactosa, D-ribosa, fructosa y glucosa (Oshodi et al., 1999). Los carbohidratos de la quinua pueden ser considerados un alimento nutracéutico porque tiene efectos hipoglicémicos beneficiosos e inducen la reducción de los ácidos grasos libres (Abugoch, 2009).

Además del alto contenido y calidad de sus proteínas, los granos de quinua presentan una interesante composición lipídica que varía entre 1.8-9.5% (Abugoch, 2009). La quinua presenta un contenido de aceites (7% peso seco) más alto que el del maíz (4.9%) y más bajo que la soya (20.9%) (Koziol, 1993). Los triglicéridos representan la mayoría de los lípidos neutrales, con un porcentaje mayor del 50%. Los diglicéridos contribuyen con el 20% de los lípidos neutrales. La lisofosfatidiletanolamina y la fosfatidilcolina son los lípidos polares totales más abundantes y representan el 57%. En cuanto a la composición de ácidos grasos de la quinua: 12.3-19% de saturados totales, principalmente ácido palmítico; 25-28.7% de monoinsaturados totales, principalmente ácido oleico; 58.3% de poliinsaturados totales, de los cuales el ácido linoleico representa el 90% de éstos (Abugoch, 2009). La fracción de aceite de los granos de quinua son de

alta calidad y altamente nutritivas, específicamente, los ácidos grasos poliinsaturados tienen efectos positivos contra las enfermedades cardiovasculares y mejoran la sensibilidad a la insulina. Otra importante fracción de lípidos la compone la vitamina E (α -tocoferol) que se encuentra en altas cantidades en la semilla (0.59-2.6 mg/100g de semilla) y actúa como una defensa natural contra la oxidación de lípidos (Coulter y Lorenz, 1990).

La quinua es también rica en micronutrientes como minerales y vitaminas. Los principales minerales son el potasio, fósforo y magnesio. El magnesio, manganeso, cobre y hierro presentes en 100 g de granos de quinua cubren las necesidades diarias de infantes y adultos, mientras que el contenido de fósforo y zinc contenido en 100 g es suficiente para niños, pero cubre 40-60% de las necesidades diarias para un adulto. El contenido de potasio contribuye entre 18% y 22% de los requerimientos para infantes y adultos, mientras que el contenido de calcio contribuye con 10% de los requerimientos (Abugoch, 2009). En cuanto a las vitaminas, los granos presentan un alto contenido total de vitamina B6 y folato, cuyas cantidades en 100 g de granos cubren los requerimientos para adultos y niños. El contenido de riboflavina en 100 g contribuye con el 80% de las necesidades diarias para niños y 40% para adultos (Ranhotra et al., 1993).

Debido a la naturaleza altamente nutritiva de la quinua, ésta es utilizada para hacer harinas, sopas, alcohol o en el desayuno. En Perú y Bolivia, se producen comercialmente hojuelas de quinua, tortillas y panqueques. La quinua puede ser vendida en forma de granos que son cocinados como el arroz, se puede fermentar para hacer cerveza o puede ser utilizada como forraje para alimentar a las vacas, cerdos y aves de corral (Galwey, 1989). La harina de quinua, en combinación con la harina de trigo o harina de maíz, es utilizada para hacer galletas, pan y comida procesada. La harina de los granos tiene una buena propiedad de gelificación, capacidad de absorción de agua, capacidad de emulsión y estabilidad. Por otro lado, la utilización de la quinua con fines medicinales se ha reportado en menor medida: la quinua se ha utilizado en inflamaciones, como analgésico y desinfectante del tracto urinario, también en fracturas,

hemorragias internas y como repelente de insectos (Mujica, 1994). La quinua también podría ser efectivamente utilizada en la industria de bebidas para la preparación de formulaciones de bebidas malteadas (Bhargava et al., 2006b). Con respecto al almidón de la quinua, éste tiene un potencial no alimenticio en su utilización como cargas biodegradables en películas de polietileno de baja densidad y en la manufacturación de bolsas donde la resistencia a la tensión es importante. Incluso las saponinas, que representan el componente no deseado en términos alimenticios, poseen una importancia industrial inmensa y son usadas en la preparación de jabones, detergentes, champús, cerveza, extinguidores, en fotografía, cosméticos y en la industria farmacéutica (Johnson y Ward, 1993).

2.2. Marcadores moleculares microsatélites

2.2.1. Composición y distribución de los microsatélites

Los microsatélites son secuencias de ADN con unidades de repetición de 2 a 6 pb de longitud, son también llamados secuencias simples repetidas (en inglés *simple sequence repeats*, SSR) o secuencias cortas repetidas (*short tandem repeats*, STR), y están presentes en todos los genomas de eucariotas y procariotas analizados hasta la fecha (Zane et al., 2002). Los patrones más abundantes en los genomas de las plantas son $(AT)_n$, $(GA)_n$ y $(GAA)_n$ donde n se refiere al número total de repeticiones y usualmente n va de 10 a 100 repeticiones (Madesis et al., 2013).

Los microsatélites están presentes tanto en regiones codificantes como no codificantes y están distribuidos a lo largo del genoma nuclear, aunque también se pueden encontrar en el genoma mitocondrial y cloroplastidial (Kalia et al., 2011). Se caracterizan por un bajo grado de repetición por locus (5-100), distribución al azar de 10^4 - 10^5 por genoma (Tautz, 1993) y un alto grado de polimorfismo (Zane et al., 2002).

Los microsatélites están generalmente asociados a otras repeticiones genómicas, especialmente elementos transponibles. En humanos, los microsatélites están asociados a ADN repetitivo, especialmente retrotransposones diferentes de los

transposones LTR (repetición terminal larga). Sin embargo, en plantas, los microsatélites están localizados preferentemente en regiones de ADN no repetitivo, lo cual indica que se encuentran en regiones previas a la expansión del genoma (Madesis et al., 2013).

Los microsatélites están uniformemente espaciados en el genoma, aunque son muy variables en número de unidades repetidas entre los individuos (Figura 2.8). Las secuencias que flanquean a los microsatélites son conservadas, por lo que se pueden desarrollar cebadores específicos para usarlos en una reacción de PCR, de esta manera los productos amplificados por el PCR pueden ser separados en sistemas de electroforesis de alta resolución (Figura 2.8) y las bandas resultantes pueden ser reveladas mediante marcaje fluorescente o tinción con plata (Jiang, 2013).

El polimorfismo de los microsatélites se debe al fenómeno deslizamiento (*slippage*, Figura 2.9) durante la replicación del ADN el cual consiste en el deslizamiento de la cadena nueva o la cadena patrón de ADN sobre la otra en al menos un motivo lo que resulta en un mal apareamiento de las cadenas (Yañez, 2002).

Los microsatélites fueron inicialmente descritos en humanos (Litt y Luty, 1989), y posteriormente se fueron descubriendo en otros organismos. Su potencial como marcadores útiles para los estudios en plantas fue rápidamente reconocido lo cual se ve reflejado en la cantidad de microsatélites obtenidos para diversas especies vegetales como el arroz, la papa, la cebada, entre otros (Yañez, 2002).

2.2.2. Clasificación de los microsatélites

Los marcadores SSR han sido clasificados de manera variada según su tamaño, tipo de unidad de repetición y localización en el genoma. Según el número de nucleótidos por unidad de repetición los microsatélites pueden dividirse en mono-, di-, tri-, tetra-, penta- o hexanucléotido. Entre los dos tipos de repeticiones mononucleótido, A/T es la más abundante en todas las especies de plantas, mientras que G/C es más limitado en número. Tanto para motivos mononucleótido y dinucleótido, los motivos ricos en CG son

menos abundantes en genomas de monocotiledóneas y dicotiledóneas. Sin embargo, para repeticiones trinucleótido, las repeticiones AGC/CGT, AGG/CCT y CCG/CGG son más frecuente en monocotiledóneas, mientras que repeticiones ricas en AT tales como AAC/GTT, AAG/CTT y AAT/ATT son más frecuentes en dicotiledóneas. La frecuencia de repeticiones tetra-, penta y hexanucleótido es baja en todos los genomas de plantas (Madesis et al., 2013).

Según la disposición de los nucleótidos en los motivos repetidos, Weber utilizó los términos perfecto, imperfecto y compuesto para su clasificación (Weber, 1990), mientras que Wang y colaboradores clasificaron a los microsatélites como perfecto simple, imperfecto simple, perfecto compuesto e imperfecto compuesto (Wang et al., 2009). Las repeticiones perfectas son arreglos en tándem de un único motivo de repetición, mientras que las repeticiones imperfectas son repeticiones perfectas interrumpidas por motivos no repetitivos en algunas partes de la región microsatélite. En los microsatélites compuestos, dos motivos repetidos básicos son presentados juntos en varias configuraciones. Jarne y Lagoda acuñaron los términos puro e interrumpido para las repeticiones perfectas o imperfectas, respectivamente (Jarne y Lagoda, 1996). Según su localización, los microsatélites pueden ser clasificados como nucleares (nuSSR), mitocondriales (mtSSR) y cloroplastidiales (cpSSR).

2.2.3. Ventajas y limitaciones de los marcadores microsatélites

Los marcadores microsatélites son de especial interés para los investigadores porque son unos de los pocos marcadores moleculares que permiten a los investigadores comprender las cuestiones ecológicas de una manera específica, además presentan ciertas características que son ventajosas con respecto a otros marcadores moleculares:

- Las técnicas basadas en ADN como los microsatélites utilizan la PCR para amplificar los marcadores de interés a partir de muestras diminutas de tejido. Debido a que los microsatélites son usualmente más cortos en longitud que los

loci secuenciados (100-300 pb vs. 500-1500 pb), estos pueden ser amplificados mediante una PCR aún si existe ADN degradado en la muestra. Esta característica permite que los microsatélites sean usados con muestras de ADN antiguas o muestras de ADN provenientes de cabello y heces (Selkoe y Toonen, 2006). Además, ya que los microsatélites son específicos para cada especie, la contaminación cruzada por organismos diferentes del organismo blanco no representa un problema.

- Aparte de la naturaleza codominante y el multiallelismo, los marcadores SSR son los que poseen el más elevado contenido de información polimórfica (PIC) de los marcadores moleculares, y aunque las técnicas de AFLP, alozimas y RAPD son técnicas multilocus, ninguna de estas tiene el poder y resolución de un estudio multilocus utilizando microsatélites. Gerber y colaboradores demostraron que 159 loci AFLP presentan menor poder estadístico para determinar paternidad que 6 marcadores microsatélites polimórficos (Gerber et al., 2000).
- Los microsatélites tienen generalmente una alta tasa de mutación que resulta en una alta diversidad alélica. Esta característica resulta útil cuando, por ejemplo, se está interesado en entender la demografía actual y los patrones de conectividad, o detectar cambios en el pasado reciente (10-100 generaciones), resolver problemas relacionados con la paternidad o estructura clonal (Selkoe y Toonen, 2006).
- Los microsatélites son muy frecuentes y están distribuidos al azar, permitiendo la más completa cobertura de cualquier genoma eucarionte (Ferreira y Grattapaglia, 1998).

A pesar de presentar muchas ventajas, los marcadores microsatélites presentan varias limitaciones que pueden complicar el análisis de los datos, y en el peor de los casos limitar su utilidad. Afortunadamente, muchas de las limitaciones comunes de los marcadores SSR pueden ser evitadas siguiendo una cuidadosa selección de los loci.

- Un par de marcadores microsatélites convencional raramente amplifica para un grupo taxonómico amplio, por lo que nuevos cebadores son sintetizados para cada nueva especie. Y aunque el proceso de aislamiento de nuevos marcadores SSR se ha vuelto más rápido y más económico, existen algunos taxones en donde el proceso de aislamiento de microsatélites presenta una alta tasa de error, como el caso de invertebrados marinos, lepidópteros y aves (Selkoe y Toonen, 2006).
- Los procesos mutacionales de las regiones microsatélites son muy complejas y para la mayoría de aplicaciones no es importante saber exactamente el mecanismo mutacional de cada locus, sin embargo, varios estadísticos basados en frecuencias alélicas (por ejemplo, F_{ST} y R_{ST}) necesitan explícitamente de un modelo mutacional. El modelo de los alelos infinitos (IAM) es el modelo más simple, y aunque exista un modelo específico para microsatélites, como el modelo mutacional por pasos (SMM), el IAM sigue siendo usado por su robustez y confiabilidad, además de que no es tan sensible a violaciones a este modelo (Selkoe y Toonen, 2006).
- La identificación de diferentes alelos de los microsatélites se basa en la diferencia del tamaño de éstos, pero existen alelos del mismo tamaño que no provienen de un antecesor común, un fenómeno denominado homoplasia. La homoplasia puede limitar la diversidad alélica visible de las poblaciones y puede inflar las estimaciones del flujo génico cuando la tasa de mutación es alta. En general, la homoplasia es frecuentemente una fuente mínima de errores para los estudios de genética de poblaciones cuya historia es reciente o con un tamaño poblacional efectivo moderado debido a que la probabilidad de que exista homoplasia es directamente proporcional a la distancia genética entre dos individuos o poblaciones, es por esta razón que la homoplasia puede representar un problema cuando se estudian grupos altamente divergentes, loci de alta tasa

de mutación o en estudios de reconstrucción filogenética (Selkoe y Toonen, 2006).

- Las regiones de unión a cebadores para la amplificación de microsatélites deben ser regiones conservadas. Si ocurren mutaciones en estas regiones, algunos individuos tendrán solo un alelo o ningún alelo amplificado, por lo que algunos loci no podrían utilizarse debido a problemas de amplificación (Selkoe y Toonen, 2006). A estos alelos que no pueden ser amplificados se les denomina alelos nulos (Wang et al., 2009).
- La mayor limitación de la tecnología de microsatélites es la gran cantidad de trabajo necesario para el desarrollo previo de los marcadores (Ferreira y Grattapaglia, 1998). Este trabajo requiere de una inversión de recursos económicos y la experiencia técnica requerida para el clonamiento y secuenciación de los loci SSR (Yañez, 2002).
- Durante la amplificación de las regiones microsatélites, el deslizamiento de la polimerasa puede ocurrir debido a las repeticiones de los SSR. Cuando los productos de PCR son separados por electroforesis, picos adicionales pueden ser observados. Estos picos adicionales se denominan picos tartamudos y son artefactos del PCR o productos no específicos usualmente de unas repeticiones más cortas de longitud, de menor cantidad y de intensidad más débil que los verdaderos alelos microsatélites. Algunos picos tartamudos interpretados como verdaderos alelos pueden afectar los análisis, sin embargo, estos picos tartamudos también pueden ser de gran ayuda para diferenciar a verdaderos alelos de otros artefactos (Wang et al., 2009).
- Los amplificados con una adenina de más (picos adenilados) también pueden confundir la identificación de los verdaderos alelos (Butler, 2005). La ocurrencia de estos artefactos dificulta la lectura de los electroferogramas porque los artefactos pueden ser confundidos con alelos, lo cual complica el análisis genotípico.

Sin embargo, estas consideraciones no han desalentado a los investigadores quienes han convertido a los microsatélites en unas herramientas moleculares muy populares (Yañez, 2002).

2.2.4. Aplicaciones de los microsatélites

Los microsatélites se han convertido en un marcador de elección para una variedad de aplicaciones en plantas debido a su naturaleza hipervariable y su extensa cobertura en el genoma. A continuación se citan algunos ejemplos de su uso:

- Son herramientas poderosas para la estimación de la diversidad genética y relaciones filogenéticas de las especies debido a su alta eficiencia, naturaleza codominante, reproducibilidad y alto grado de polimorfismo. Esta información será importante para escoger líneas parentales para los programas de mejoramiento, clasificación de las accesiones de los bancos de germoplasma y posterior curación y adquisición de nuevas accesiones para los bancos (Kalia et al., 2011).
- Para la identificación de cultivares y el análisis de pedigree ya que cada microsatélite representa un solo locus.
- En la determinación de la hibridación, donde la naturaleza codominante de los microsatélites juega un papel importante y permite detectar la contribución alélica de cada padre en híbridos somáticos o sexuales.
- En la determinación del sexo de plantas sexualmente dimórficas, como en el caso de la papaya (Parasnis et al., 1999), los marcadores microsatélites ligados al sexo se utilizan para seleccionar las plantas hembras, que son valiosas comercialmente, antes de haber alcanzado la madurez y diferenciarse por completo de la planta macho (estadío de plántula).
- Los marcadores microsatélites específicos de organelas (mtSSR y cpSSR) han tenido un gran impacto en la determinación de la estructura y variación dentro de una población natural, así como las relaciones filogenéticas.

- En el análisis de loci de carácter cuantitativo (QTL) lo cual conlleva a la identificación de genes candidatos para los rasgos de interés que son particularmente importantes para los programas de mejoramiento tales como el rendimiento, resistencia a enfermedades, tolerancia al estrés, calidad de la semilla y el fruto, etc (Kalia et al., 2011), y el posterior diseño de genotipos superiores.
- Los microsatélites basados en marcadores de secuencia expresada (EST-SSRs) pueden contribuir a la selección alélica directa debido a que éstos tienen funciones putativas o conocidas y pueden ser asociados a rasgos blanco (Kalia et al., 2011).
- Los microsatélites también han sido extensamente utilizados en el mapeo genómico que consiste en el mapeo genético, comparativo, físico y de asociación (Kalia et al., 2011).

2.2.5. Trabajos previos en quinua utilizando microsatélites

Un equipo de investigación de la universidad Brigham Young en los Estados Unidos desarrolló una colección de aproximadamente 450 marcadores microsatélites para *C. quinoa* (Maughan et al., 2004; Mason et al., 2005; Jarvis et al., 2008). Los marcadores SSR reportados han sido utilizados exitosamente para evaluar la diversidad genética en diferentes colecciones de germoplasma y para la construcción de mapas genéticos de ligamiento.

Mason y colaboradores (2005) desarrollaron a gran escala los primeros marcadores microsatélites para quinua que consistieron de 208 marcadores polimórficos los cuales fueron validados y caracterizados utilizando una colección de 31 accesiones de quinua provenientes de las principales regiones de cultivo en Sudamérica. En total, 1276 clones fueron secuenciados a partir de 3 librerías enriquecidas de motivos microsatélites (CA, AAT, ATG), de los cuales 407 clones (36%) contenían microsatélites únicos. Los motivos de repetición más comunes, aparte de CA, AAT y ATG, fueron los motivos GA y CAA.

Dentro de los motivos más comunes, el motivo AAT es más prevalente en el genoma de la quinua. Se diseñaron 397 cebadores que flaqueaban los 397 loci microsatélites y se probaron en un panel de diferentes accesiones de *C. quinoa* y una accesión de *C. berlandieri* Moq., un relativo silvestre de la quinua. 208 marcadores microsatélites (52%) fueron polimórficos entre las accesiones de quinua. 25 marcadores adicionales (6%) eran polimórficos cuando la accesión de *C. berlandieri* era incluida en el análisis.

El análisis genético realizado en la colección de quinua reveló un número observado de alelos que variaba de 2 a 13, con un promedio de 4 alelos detectados por locus. Los valores de heterocigocidad variaron de 0.2 a 0.9 con un valor promedio de 0.57. 67 marcadores (32%) fueron altamente polimórficos ($H \geq 0.7$). Este set de marcadores microsatélites reveló un potencial de utilidad para estudios moleculares utilizando especies relacionadas de la subfamilia Chenopodioidae.

La amplificación de 202 de los 208 marcadores SSR polimórficos en un grupo que consistía de 2 accesiones de *C. berlandieri* subsp. *nutalliae* (huauzontle), 2 de *C. pallidicaule* (cañihua) y 2 de *C. giganteum* reveló que 67% de los marcadores amplificaron exitosamente para todos los grupos. El más alto nivel de conservación en amplificación por PCR fue observado en *C. berlandieri* subsp. *nutalliae*, con 99.5% de amplificaciones reproducibles, de las cuales 81% fueron polimórficas sobre la base de las dos accesiones utilizadas, lo cual confirma el parentesco cercano con *C. quinoa*.

Christensen et al. (2007) utilizaron 32 SSR identificados por Mason et al. (2005) y 4 identificados por Maughan et al. (datos no publicados) para evaluar la diversidad genética de 152 accesiones de quinua de las colecciones de los bancos de germoplasma del departamento de agricultura de los Estados Unidos (USDA) y del centro internacional de la papa (CIP). Se detectaron 420 alelos, con un rango de 3 a 27 alelos por locus y con un promedio de 11 alelos por locus. También reportaron una heterogeneidad genética en el 32% de las accesiones en un locus dado lo cual sugirió que muchas de las accesiones representaban lotes heterogéneos de semillas o variedades nativas.

A partir de un análisis por UPGMA y análisis de componentes principales se dividió a las accesiones en dos grandes grupos: el primer grupo consistió de las accesiones del altiplano de Perú, Bolivia, Ecuador, Argentina, y del noreste de Chile; y el segundo grupo incluyó a las accesiones de la zona costera de Chile y las accesiones cuya información de procedencia se desconocía (Christensen et al., 2007).

Jarvis y colaboradores (2008) desarrollaron un nuevo conjunto de marcadores SSR y construyeron un mapa genético de ligamiento para quinua. Este trabajo describe el desarrollo de 216 nuevos marcadores microsatélites. Las repeticiones observadas más comunes en este estudio fueron GA, CAA y AAT. De igual manera, los motivos CA, CGA, GAA y GGT también fueron observados, aunque en menor frecuencia. La evaluación de estos 216 marcadores SSR en 22 accesiones de quinua representativas de los andes en Sudamérica reveló un número de alelos observados por marcador que varió de 2 a 13, con un promedio de 4 alelos por marcador con valores H que variaron de 0.12 a 0.9, con un promedio de 0.57, valores similares a aquellos reportados por Mason y colaboradores (2005). Los 216 marcadores microsatélites fueron considerados polimórficos ($H \geq 0.7$). En adición, las diferencias polimórficas entre las repeticiones dinucleótidas y trinucleótidas confirmaron la observación común de que el desarrollo de marcadores microsatélites altamente polimórficos debería estar proyectado para la utilización de microsatélites trinucleótidos con longitudes mayores a 20 pb.

Fuentes y colaboradores (2009a) caracterizaron y cuantificaron la diversidad genética de 28 accesiones de quinua del altiplano de Chile y 31 accesiones de la costa de Chile utilizando 20 marcadores SSR di- y trinucleótidos desarrollados por los dos únicos estudios anteriores (Mason et al., 2005; Jarvis et al., 2008), con los cuales se desarrollaron 7 grupos de marcadores fluorescentes para un sistema multiplex.

La selección de los 20 marcadores SSR se basó en los niveles de heterocigosidad, reproducibilidad y calidad de amplificación. Un total de 150 alelos fueron detectados entre las accesiones en un rango de 2 a 20 alelos por locus y un promedio de 7.5 alelos por locus. Los análisis por agrupamiento (UPGMA) y análisis de componentes

principales separó las accesiones en dos grupos discretos similar a los hallazgos en los análisis de isozimas y rasgos morfológicos (Wilson, 1988), análisis de AFLPs (Pratt, 2003) y microsatélites (Christensen et al., 2007). El primer grupo contenía accesiones de quinua del norte (altiplano) y el segundo grupo consistía de accesiones provenientes del sur (costa). Las quinuas costeras presentaron una mayor diversidad que las del altiplano, con un 50% de alelos únicos en las accesiones costeras y un 28.6% en las del altiplano, y ambos grupos compartieron un 21.3% de alelos. Se concluyó que existía una diversidad reducida en las accesiones del norte de Chile y se reportó un nuevo subconjunto de marcadores genéticos altamente informativos y fácil de utilizar (Fuentes et al., 2009a).

González (2009) estudió el flujo de genes en quinua cultivada, maleza y silvestre de la sierra de Ecuador utilizando marcadores microsatélites, seleccionó 8 marcadores microsatélites por su alto polimorfismo y utilizó el método de la cola M13 en un secuenciador LI-COR 4300S para evaluar la variabilidad genética de 46 individuos provenientes de los campos de producción (de grano blanco y maleza conocida como “malla”), 10 individuos de la variedad INIAP-Tukauan y 21 individuos silvestres. Un total de 84 alelos fueron reportados para los 77 individuos evaluados cuyos pesos oscilaron entre 165 a 199 pb. El locus que presentó mayor variabilidad fue el locus QCA26. Se determinó que existía un problema de mezcla en las panojas de grano blanco con grano negro, se debía a la existencia de flujo genético entre las poblaciones cultivadas y sus parientes silvestres, generando así plantas híbridas. Se reveló una afinidad genética entre los grupos de quinua-maleza y sus parientes silvestres, y una clara separación entre las poblaciones cultivadas y silvestres.

Costa y colaboradores (2012) colectaron 35 accesiones de quinua del noroeste argentino para estudiar su diversidad y estructura genética utilizando 22 marcadores microsatélites. Sus resultados mostraron un gran nivel de diversidad genética, el número total de alelos detectados en las 35 poblaciones de quinua para los 22 loci SSR fue de 364, con un número promedio de alelos por locus de 16, con valores que iban desde 6

(QCA067) hasta 29 (QAAT022). Todos los loci fueron altamente polimórficos con valores de heterocigocidad (H_e) mayores que 0.7 a excepción del locus QAAT084 el cual presentó un valor de 0.58. Se agruparon las accesiones en 4 principales grupos en una distancia genética promedio moderada (0.8), cada una de las cuales representa un diferente ambiente de la región del noroeste argentino: Puna (C2), valles secos (C3), valles húmedos del este (C4) y un área de transición (C1) entre los dos últimos ambientes; con unos valores de U_{He} de 0.42, 0.27, 0.16 y 0.25 respectivamente. Los análisis por AMOVA mostraron una fuerte estructuración genética: una alta subdivisión poblacional relativa al agrupamiento por regiones ($F_{SR}=0.47$) junto con una alta diferenciación genética entre poblaciones ($F_{ST}=0.58$) y un número bajo de heterocigotos ($F_{IS}=0.63$) en cada uno de ellos. Se observó una gradiente decreciente de la diversidad genética (dirección oeste a este) siendo la Puna (C2) el grupo más diverso orientado hacia el oeste y los valles húmedos del este (C4) el menos diverso y orientado hacia el este. Este trabajo fue la primera caracterización molecular de una muestra representativa del germoplasma de quinua de Argentina. Es impresionante que una alta diversidad haya sido encontrada a pesar que el noroeste de Argentina representa una pequeña área geográfica.

Fuentes et al. (2012) reportaron el impacto del intercambio de semillas entre agricultores y las prácticas de producción local sobre la estructura y diversidad genética de quinua a escala nacional en Chile. Utilizaron 20 marcadores microsatélites en un conjunto de 34 accesiones de quinua representativa de Chile y de la región de Sudamérica. El valor H para todas las accesiones de quinua variaron entre los valores 0.12 (QGA17) y 0.87 (QAAT76) y un valor promedio de 0.65. El análisis por UPGMA identificó dos grandes grupos las cuales fueron subdivididas en 5 poblaciones: La población I contenía 9 accesiones representativas de la zona norte de Chile; la población II, 7 accesiones de la zona central; la III, 9 accesiones de la zona sur y solo una de la zona central; la IV, 6 accesiones del altiplano del Perú, Bolivia y Argentina; y la V, 2 accesiones de Ecuador y Colombia. El origen geográfico de las accesiones y las poblaciones identificadas (I-V)

fueron consistentes con las clasificaciones de ecotipos (Risi y Galwey, 1984): La población I fue representativa de los salares (norte de Chile), la II y la III de costa del sur de Chile, la IV del altiplano (Perú, Bolivia y Argentina), y la V de los valles interandinos del norte de Sudamérica (Ecuador y Colombia). Las poblaciones III (costa de Chile) y IV (altiplano del Perú, Bolivia y Argentina) presentaron los valores más altos de H y proporción de loci polimórfico, mientras que la población III tuvo los valores más altos para el número promedio de alelos por locus y número de alelos totales a pesar de tener una menor dispersión geográfica que la población IV. La población I exhibió el valor máximo de alelos únicos, seguido por las poblaciones III y IV, y las poblaciones II y V rindieron valores bajos para todas las variables que describen el patrón de diversidad alélica.

Veramendi y colaboradores (2013) seleccionaron cerca del 90% de la colección boliviana de quinua compuesta de materiales provenientes del altiplano boliviano (centro con 827, norte con 138 y sur con 434 accesiones respectivamente), de los valles (325 accesiones), del Perú (562 accesiones), de materiales silvestres (133 accesiones) y una fracción proveniente de otros países (96 accesiones) para caracterizar la diversidad genética utilizando 8 marcadores microsatélites. Se detectaron un total de 129 alelos entre las diferentes regiones analizadas, en un rango de 5 (QCA006) y 30 (QAAT022) con un promedio de 16 alelos por locus, que van de 111 a 239 pb. El PIC para la colección total presentó valores entre 0.73 y 0.95 con un promedio de 0.84, resultando todos los marcadores altamente polimórficos. Los marcadores QAAT074, QAAT076 y QAAT022 resultaron ser los más polimórficos. El dendrograma, obtenido mediante el método de agrupamiento UPGMA y coeficiente de similitud de Jaccard permitió la formación de grupos con accesiones en su mayoría de la misma región. También se pudo elaborar una colección núcleo de quinua con 189 accesiones representativas de la colección total.

Costa (2014) determinó la magnitud de la diversidad genética y estructuración genética de 36 poblaciones (accesiones porque no se utilizó la técnica del bulk) del germoplasma

del noroeste de Argentina utilizando 22 loci microsatélites. El valor de heterocigocidad fue de 0.3 y se detectaron 360 alelos en todas las poblaciones, de los cuales 97 fueron alelos únicos. El promedio de alelos detectados por locus fue de 16, aproximadamente en un rango de 6 (QCA037) a 33 (QAAT022). Considerando los valores de heterocigocidad total, la mayoría de los loci SSR caracterizados fueron altamente polimórficos, presentando valores superiores a 0.7 a excepción de los loci QAAT084 y QGA002 (0.58 y 0.62 respectivamente). El rango de los valores de riqueza alélica (número de alelos) entre las poblaciones fue entre 20 y 124 con un promedio de 50 ± 4.1 . La media aritmética de U_{He} entre todas las poblaciones fue 0.29 ± 0.01 , en un rango entre 0 y 0.71 ± 0.03 . Los valores extremos de ambos parámetros fueron para las poblaciones CHEN274 y 272 respectivamente. El promedio de loci polimórficos entre todas las poblaciones fue 62.1 ± 4.9 . La población CHEN274 no presentó polimorfismo en ningún loci, mientras que las poblaciones CHEN269, 272, 432 y 427 presentaron polimorfismo en el 100% de los loci analizados, siendo las ultimas 4 las poblaciones con mayor riqueza alélica, mayor diversidad genética de la colección explorada y todas procedentes de la puna. Por medio del AMOVA se determinó que el 18% del total de la varianza se debió a la diferenciación entre regiones ($F_{RT}=0.18$), 39% entre poblaciones ($F_{ST}=0.57$), 27% entre plantas individuales y el 16% restante intra-individuos. Los valores de F_{IS} (0.63) y F_{IT} (0.84) indicaron una deficiencia de genotipos heterocigotos. El agrupamiento obtenido por UPGMA reveló una estructuración del germoplasma de manera longitudinal en la región del noroeste, formándose 4 grandes grupos de poblaciones que correspondían a regiones agroecológicamente diferentes: Puna (G4), valles secos (G1), valles húmedos (G2) y una zona de transición de altura (G3). De oeste a este, la magnitud de la diversidad genética a nivel de grupo presentó un gradiente decreciente. Las poblaciones procedentes de la puna (12) fueron en promedio las más diversas ($U_{He}=0.42 \pm 0.02$), el 77.65% de los loci caracterizados, en promedio, fueron polimórficos amplificándose en total 797 alelos de los cuales 41 fueron exclusivos.

Vía y Rada (2015), con el fin de describir la diversidad genética y elucidar la estructura poblacional de la quinua en el Perú, estimó la diversidad genética de 172 accesiones de quinua procedentes de valles interandinos y del altiplano mediante la genotipificación utilizando 23 marcadores microsatélites mediante un sistema de PCR multiplex. Se detectaron 294 alelos en total con un promedio de 12.78 alelos por locus, siendo los ecotipos de valles interandinos los que presentaron un mayor número de alelos exclusivos (60 alelos), por lo tanto, esta población presentó mayor riqueza alélica. El marcador QAAT100 presentó el mayor número de alelos (44 alelos), mientras que el marcador KCAA065 presentó un único alelo, siendo este último un locus monomórfico. El 20% de los alelos detectados fue exclusivo de quinuas de valles interandinos, el 15% fue exclusivo de quinuas del altiplano y un 65% del total de alelos fue compartido por ambos grupos. Se realizó un análisis de coordenadas principales en donde se observó la división poblacional entre las accesiones del altiplano y las de los valles interandinos. Se logró identificar dos subpoblaciones de quinua con diferenciación genética moderada ($F_{ST}=0.059$), las cuales guardaron relación con la procedencia de las muestras. De igual manera, se realizó un análisis discriminante de componentes principales (DAPC) para identificar grupos genéticos y se determinó que existe una contribución a más de un grupo genético en cada individuo (159 individuos con probabilidades de pertenencias a más de un grupo), además que no existió una estructuración genética definida dentro y entre los grupos inferidos por el DAPC. También se pudo identificar 10 marcadores altamente polimórficos (KGA3, QAAT106, QAAT24, QAAT50, QAAT70, KAAT037, QGA024, QAAT76, QAAT74 y QAAT100) recomendados para su utilización en la evaluación de la diversidad genética del germoplasma de quinua del banco de germoplasma del Instituto Nacional de Innovación Agraria (INIA).

2.3. Métodos de separación, detección y marcaje del ADN

2.3.1. Electroforesis capilar

La electroforesis capilar es una técnica de separación basada en la diferente velocidad de migración de las distintas moléculas cargadas bajo la acción de un campo eléctrico (Osatinsky, 2007). Es una nueva adición a la familia de electroforesis, las primeras separaciones de ADN fueron realizadas en los últimos años de la década de los 80 (Butler, 2005).

Los principales instrumentos necesarios para realizar una electroforesis capilar incluyen capilares angostos, dos viales con soluciones tampón y dos electrodos conectados a una fuente de poder de alto voltaje (Figura 2.10). Los sistemas de electroforesis capilar también contienen una fuente de excitación laser, un detector de fluorescencia, un muestreador automático para sostener la placa con las muestras y una computadora para controlar la inyección y detección de la muestra. (Butler, 2005).

Los capilares están hechos de sílice fundida cuyos diámetros varían de 50-100 μm y una longitud de 25-75 cm. Una solución de polímero viscosa (POP) recorrerá los capilares y será el medio por donde pasarán las muestras. Los campos eléctricos que se utilizan en la electroforesis capilar son 10 a 100 veces más fuertes que los campos utilizados en la electroforesis convencional (300 V/cm en vez de 10 V/cm), lo que permite una mayor velocidad de corrida. La detección de la muestra es realizada automáticamente por el instrumento mediante la medición del intervalo de tiempo desde la inyección de la muestra hasta la detección de la muestra con un láser colocado cerca del extremo del capilar. El láser de gas de iones de argón, que produce luz a 488 nm y 514.5 nm, ilumina en una parte del capilar donde se encuentra una ventana por donde la luz entra en contacto con las muestras que atraviesan el capilar. El detector de la fluorescencia es un aparato fotosensible que mide la intensidad de luz emitida por los fluoróforos y consiste en un dispositivo con carga acoplada (CCD). Los datos obtenidos de la electroforesis se representan como una función de la intensidad de fluorescencia relativa (unidades de fluorescencia relativa, RFU) observada desde la emisión de

fluorescencia de fluoróforos que pasan por el detector. Las señales de emisión de fluorescencia de los fluoróforos unidos a las moléculas de ADN son usadas para detectar y cuantificar las moléculas de ADN que pasan a través del detector. Las muestras se corren junto con unos marcadores estándares marcados con un fluoróforo de distinto color que el de las muestras para calibrar los tiempos de migración de los fragmentos de ADN de interés con las muestras de tamaño conocidas (Butler, 2005).

Algunas de las ventajas de la electroforesis capilar sobre la electroforesis convencional son:

- Las etapas de inyección, separación y detección pueden ser totalmente automatizadas permitiendo que múltiples muestras sean procesadas sin ninguna supervisión.
- Se consumen cantidades ínfimas de muestra en la etapa de inyección, también pueden ser reutilizadas si se necesitan.
- Las separaciones en capilares pueden ser conducidas en minutos debido a los altos voltajes que son permitidos con una mejorada disipación del calor.
- La información cuantitativa se encuentra disponible en un formato electrónico después de acabada la corrida.
- No hay peligro de contaminación cruzada entre diferentes muestras debido a que cada muestra está contenida en cada capilar.

Por otro lado, una de las desventajas principales de la electroforesis capilar es que los instrumentos para llevarla a cabo requieren costos iniciales elevados, llegando a costos de más de \$50000.

2.3.2. Detección y marcaje del ADN por fluorescencia

La detección de moléculas de ADN basada en fluorescencia involucra el uso de fluoróforos excitados y la detección de la luz emitida por estos. Este tipo de detección es utilizada ampliamente por su capacidad de análisis multicolor, así como su facilidad de uso y el corto tiempo que se emplea para la detección (Butler, 2005).

La elección del color del fluoróforo depende del criterio del investigador, pero debe ser consistente con los filtros instalados del respectivo secuenciador ABI. Se prefiere el fluoróforo 6-FAM (520 nm de emisión máxima aproximada o AME, color azul a verde) con respecto a los otros fluoróforos debido a su fluorescencia adecuada y que no es propiedad de ninguna compañía y por lo tanto es más barato. Para el analizador genético ABI 3730xl, se utilizan típicamente los fluoróforos VIC® (553 nm de AME, color verde), NED® (575 nm de AME, color amarillo) y PET® (590 nm de AME, color rojo) que son propiedad única de Applied Biosystems™. Los fluoróforos LIZ (650 nm de AME, color naranja) y ROX (608 nm de AME, color naranja) se encuentran reservados únicamente para los marcadores estándar de tamaño (Culley et al., 2013).

El marcaje por fluorescencia de productos de PCR puede ser llevado a cabo de tres distintas maneras: (1) Incorporando un fluoróforo en el producto amplificado a través de un cebador oligonucleótido marcado en el extremo 5' (*"dual primer labeling"* o marcaje dual de cebadores, Figura 2.11); (2) incorporando desoxirribonucleótidos marcados fluorescentemente en el producto de PCR; y (3) utilizando un fluoróforo intercalador para unirse al ADN. En el análisis de regiones microsatélites amplificadas, el fluoróforo es unido al cebador utilizado en el PCR que luego es incorporado en la región diana del ADN (marcaje dual de cebadores), luego los diferentes alelos son visualizados como picos en un electroferograma (Culley et al., 2013).

El marcaje dual de cebadores es realizado por un triplete de cebadores: un cebador directo específico para un locus con la secuencia de un cebador universal en el extremo 5', un cebador reverso específico para un locus y un cebador universal marcado fluorescentemente (Schuelke, 2000). Durante la PCR, el cebador directo se agota primero debido a la baja concentración inicial y permite que el cebador universal sea posteriormente incorporado (Blacket et al., 2012). Para que el cebador directo se agote y sea incorporado primero en los productos de PCR iniciales, los primeros ciclos presentan una temperatura de hibridación específica para el par de cebadores directo y reverso. Luego de haberse agotado el cebador directo, la temperatura de hibridación se

disminuye para facilitar la incorporación del cebador universal. Por lo tanto, el cebador universal releva al cebador directo y permite el marcaje fluorescente de los productos de PCR (Schuelke, 2000).

El cebador universal puede ser una secuencia única que no sea parte del genoma del organismo de estudio y que no interfiera con otros cebadores en la reacción, por ejemplo, la secuencia M13(-21) (18 pb, 5'-TGTAACGACGGCCAGT-3'), microsatélites humanos (en el estudio de microsatélites de genomas vegetales) o alteraciones de cebadores específicos de genomas (Culley et al., 2013).

La adopción del sistema de triplete de cebadores para la amplificación de regiones microsatélites puede ser una alternativa de bajo costo si se emplea más de un solo cebador universal y más de un solo fluoróforo. Los costos se pueden reducir aún más si se utiliza un sistema de PCR multiplex para que marcadores microsatélites que se solapan en longitud puedan ser coamplificados y marcados con distintos fluoróforos (Blacket et al., 2012; Culley et al., 2013). Los costos también son reducidos si varios amplificados que se solapan en tamaño y presentan diferentes fluoróforos son combinados en un mismo pocillo (*poolplexing*) y así puedan ser corridos en un solo capilar del sistema de electroforesis capilar (Guichoux et al., 2011).

2.4. La diversidad genética y su análisis

2.4.1. Diversidad Genética

La diversidad genética es la variedad de alelos y genotipos presentes en un grupo bajo estudio, ya sea una población, especies o un grupo de especies (Frankham et al., 2002). Es la variación de características hereditarias presentes en una población de una misma especie (Osawaru et al., 2015). Representa el primer nivel de la biodiversidad, seguido de la diversidad de especies y diversidad de ecosistemas (De Vicente et al., 2003). La diversidad genética reside en cambios en las secuencias de las 4 bases nitrogenadas del ADN y permite que las poblaciones se adapten a las condiciones cambiantes del ambiente (De Vicente et al., 2003).

Las poblaciones cambian o evolucionan porque sus frecuencias génicas experimentan cambios. Estos cambios en las frecuencias génicas son causados por fuerzas que moldean la diversidad genética. Estas fuerzas son las siguientes (De Vicente et al., 2003; Osawaru et al., 2015):

- **Mutación:** Es el origen de toda la nueva diversidad genética y ocurre cuando existen errores ocasionales en la replicación del ADN u otros elementos de producción y empaquetamiento de la información genética dentro de la célula. Pueden ser perjudiciales o beneficiosas. Las mutaciones incrementan la diversidad genética, pero estas ocurren a unas bajas tasas de frecuencia, por lo que éstas por si solas no conducen la evolución de especies y poblaciones.
- **Migración:** Es el movimiento de individuos o cualquier forma de introducción de genes de una población a otra. La migración no solo implica el movimiento de individuos dentro de nuevas poblaciones, sino también la introducción de nuevos alelos dentro de las poblaciones (flujo génico). El efecto inmediato de la migración es el aumento de la variabilidad genética poblacional, causando cambios significativos en las frecuencias génicas.
- **Recombinación:** Es el proceso mediante el cual una célula genera nuevas combinaciones cromosómicas en comparación a la de sus progenitores. Genera diversidad alélica (recombinación intragénica) y diversidad genómica (nuevas combinaciones multigénicas).
- **Selección:** Es quizá uno de los procesos más conocidos que afectan a la diversidad genética y el único que resulta directamente en poblaciones mejor adaptadas al ambiente. Para que la selección natural ocurra, debe haber diferencias en la capacidad de supervivencia de los individuos a las condiciones adversas y una base genética para tales diferencias. Con el tiempo, aquellos individuos que estén mejor adaptados al ambiente vivirán más y producirán descendencia que heredaran los mismos rasgos adaptativos.

- **Deriva genética:** Se refiere a las fluctuaciones en las frecuencias alélicas de manera aleatoria (particularmente en poblaciones pequeñas) como resultado del muestreo aleatorio entre gametos. La deriva decrece la diversidad dentro de una población porque tiende a eliminar los alelos raros, reduciendo el número total de alelos.

2.4.2. Análisis de la diversidad genética

El análisis de la diversidad genética no solo consiste en cuantificarla intrapoblacional e interpoblacionalmente, sino también consiste en el cálculo de las relaciones entre las unidades (individuos) mediante el uso de distancias genéticas y expresar las relaciones mediante ciertos métodos. De esta manera, De Vicente y colaboradores (2003) dividen el análisis de la diversidad genética en tres pasos:

- (1) Descripción de la diversidad genética de manera intrapoblacional e interpoblacional, también se extiende el análisis a áreas y regiones.
- (2) Cálculo de las relaciones entre las unidades analizadas en el paso (1). Esto comprende el cálculo de las distancias entre todos los pares de unidades del estudio.
- (3) Expresión y observación de estas relaciones con algún método de agrupamiento y/u ordenación. Algunos de estos métodos permiten comparar los resultados utilizando varios tipos de datos (por ejemplo, moleculares y geográficos) y los resultados pueden ser graficados en un plano y utilizando puntos que representen las unidades analizadas.

2.4.2.1. Descripción de la diversidad genética

Para cuantificar, describir o medir la diversidad genética intrapoblacional e interpoblacional, se utilizan una serie de medidas que reflejan la riqueza y uniformidad (frecuencia relativa) de los alelos (Hughes et al., 2008). Algunas de las medidas utilizadas son:

Algunas de las medidas para cuantificar la diversidad genética intrapoblacional son las siguientes:

- Diversidad alélica (A): Es el número promedio de alelos por locus. Se requiere contar el número de alelos por locus y luego calcular el promedio.
- Riqueza alélica: El número total de alelos.
- Número efectivo de alelos (A_e): Es el número de alelos que pueden estar presentes en una población. El parámetro mide el número de alelos que se esperaría para un locus de una población. Se calcula de la siguiente manera:

$$A_e = \frac{1}{\sum p_i^2}$$

Donde p_i es la frecuencia de cada alelo y los valores son sumados para todos los alelos.

- Heterocigosidad observada (H_o): Es el número de heterocigotos en un locus dividido por el número total de individuos muestreados.
- Heterocigosidad esperada (H_e): Es la probabilidad que, en un solo locus, dos alelos elegidos de manera aleatoria, sean diferentes entre sí (De Vicente et al., 2003). Es la proporción esperada de heterocigotos si la población se aparea al azar (Allendorf y Luikart, 2007). Nei se refirió a este parámetro como diversidad génica y de abreviatura H (Nei, 1987). Se calcula de la siguiente manera:

$$H_e = 1 - \sum p_i^2$$

Donde p_i es la frecuencia del i-ésimo alelo. Se prefiere reportar la H_e que la H_o ya que es menos afectada por el tamaño de la muestra y se requiere de pocos individuos para estimar la H_e si un gran número de loci son examinados. La H_e promedio de todos los loci es un estimado del grado de variabilidad genética en la población.

Para la cuantificación de la diversidad genética interpoblacional se utilizan los denominados estadísticos de diferenciación genética interpoblacional como el índice de fijación F_{ST} (de los estadísticos F) y otros índices de fijación análogos al F_{ST} :

- Estadísticos F: Son las medidas más antiguas y ampliamente utilizadas de endogamia total en una población dividido en dos componentes: la endogamia dentro de las subpoblaciones y la endogamia debido a la diferenciación entre las subpoblaciones (Frankham et al., 2002). Desarrollado por Sewall Wright (1950), los estadísticos describen la distribución de la variación genética dentro de las especies (Allendorf y Luikart, 2007). Para hallar los estadísticos F en el caso de loci con solo dos alelos, Wright describe 3 valores de heterocigosidad calculados de manera diferentes: H_I (heterocigosidad observada subpoblacional promedio), H_S (heterocigosidad esperada subpoblacional promedio) y H_T (heterocigosidad total, utilizando las frecuencias alélicas subpoblacionales promedio). Los estadísticos F son los siguientes:

$$F_{IS} = 1 - \frac{H_I}{H_S}$$

$$F_{IT} = 1 - \frac{H_I}{H_T}$$

$$F_{ST} = 1 - \frac{H_S}{H_T}$$

Donde F_{IS} es la proporción de la endogamia total dentro de la población debido a la endogamia entre las subpoblaciones (no confundir con el coeficiente de endogamia F que es definido como la probabilidad que dos alelos en un locus dentro de un individuo sean idénticos por descendencia), F_{IT} es la endogamia total de una población debido a la endogamia dentro de las subpoblaciones y la diferenciación entre estas subpoblaciones, y F_{ST} que es la proporción de la endogamia total de la población debido a la diferenciación entre subpoblaciones (mide el grado de diferenciación genética entre las subpoblaciones en términos de frecuencias alélicas, también llamado índice de fijación). El índice de fijación F_{ST} está entre los valores 0 a 1 donde 0 indica que todas las subpoblaciones tienen iguales frecuencias alélicas y 1 que indica que todas las subpoblaciones están fijadas para diferentes alelos (Allendorf y Luikart, 2007). Nei extendió el

análisis para tres alelos a más denominando a los estadísticos G_{IS} , G_{IT} y G_{ST} de modo análogo a los estadísticos F (Nei, 1978), es por esta razón que los estadísticos F y G son habitualmente intercambiables en la literatura.

Se han propuesto diferentes enfoques dependiendo del tamaño de la muestra, tipo de marcadores genéticos y el tipo de preguntas que se quieren responder acerca de la estructura genética poblacional. Otras medidas análogas al índice de fijación F_{ST} son, por ejemplo, θ y Φ_{ST} (Bird et al., 2011).

El AMOVA (Excoffier et al., 1992) es un método para estudiar la variación molecular dentro de las especies cuyas poblaciones se encuentran estructuradas en distintos niveles jerárquicos (por ejemplo, una población total en estudio dividida en regiones y estas regiones divididas a su vez en subpoblaciones). El estadístico Φ_{ST} , a diferencia del estadístico F_{ST} , es un índice de fijación basado en distancias genéticas (Bird et al., 2011) y se obtiene utilizando el método AMOVA que es un análogo del enfoque estadístico estándar del análisis de varianzas (ANOVA) (Allendorf y Luikart, 2007).

2.4.2.2. Cuantificación de las relaciones genéticas

Otra forma de evaluar la diferenciación entre individuos es a través de los coeficientes de disimilitud (Reif et al., 2005). La disimilitud forma parte de la familia de funciones métricas y está relacionada con la distancia euclidiana. La distancia euclidiana entre dos puntos es la longitud de la línea recta que une a estos puntos. Para m puntos, las distancias calculadas entre todos los pares de puntos son escritas en una matriz denominada la matriz de distancia de tamaño $m \times m$. Una función d_{jk} es una distancia euclidiana si cumple las siguientes condiciones (teoremas métricos) para todos los puntos:

- (1) Si los puntos j y k son iguales, entonces $d_{jk} = 0$
- (2) Si $j \neq k$, entonces $d_{jk} \geq 0$
- (3) El teorema de la simetría, $d_{jk} = d_{kj}$

- (4) El teorema de la desigualdad del triángulo que postula que dado los puntos i, j y k , se cumple la siguiente relación: $d_{ij} + d_{jk} \geq d_{ik}$, en otras palabras, la distancia entre dos puntos (j y k) no puede ser más grande que la suma de sus distancias a un tercer punto (i).

Las distancias euclidianas son preferidas con respecto a otras funciones métricas debido a que son fáciles de visualizar en espacios tridimensionales a diferencia de los espacios no métricos, y la mayoría de métodos multivariados asumen que todos los puntos se encuentran en un espacio euclidiano, tales como el análisis de coordenadas principales, análisis de agrupamiento jerárquico y clasificación jerárquica (Reif et al., 2005). La mayoría de métodos de ordinación imponen limitaciones métricas, a excepción del escalamiento multidimensional no métrico. Sin embargo, hay casos donde el investigador debe usar funciones no métricas ya que estas podrían detectar mejor la estructura de los datos (Podani, 2000).

La disimilitud es aquella función que cumple con las 3 primeras condiciones de la métrica (la desigualdad del triángulo no necesariamente se debe de cumplir). El concepto de disimilitud es más general que el concepto de distancia euclidiana y el de métrica: elevar al cuadrado a una función de distancia te da como resultado una disimilitud, por lo que ambas funciones están relacionadas y pueden ser convertidas una en otra dependiendo de qué tipo de matriz se requiera. A diferencia de la distancia euclidiana, la mayoría de las funciones de disimilitud no puede exceder un cierto valor máximo, de tal manera que la mayoría de matrices de disimilitud son construidas para que el valor máximo sea 1 (máxima disimilitud) y el valor mínimo 0 (no existe disimilitud), esto es una ventaja ya que permite la comparación entre diferentes medidas de disimilitud (Podani, 2000).

La similitud (s_{jk}) es otra medida relacionada con la distancia y disimilitud, se define como el complemento de una disimilitud de rango $[0,1]$, por lo que $s_{ij} = 1 - d_{jk}$, la función d_{jk} en esta ecuación sería una disimilitud (Podani, 2000). La relación de la similitud con la distancia puede ser lineal ($D_{jk} = 1 - s_{jk}$, donde D_{jk} es una función de disimilitud que es a

su vez una distancia euclidiana), cuadrática ($D_{jk} = \sqrt{1 - s_{jk}}$) y circular ($D_{jk} = \sqrt{1 - s_{jk}^2}$) (De Vicente et al., 2003).

Existen coeficientes de distancia que trabajan con frecuencias alélicas y que necesitan de modelos de evolución para que las distancias reflejen el tiempo transcurrido desde la divergencia de los individuos o poblaciones estudiados, a estas distancias se les denominan distancias genéticas. Aquellas distancias que no toman en cuenta modelos evolutivos solo tendrán un significado geométrico (distancias geométricas) y no servirán de base para una explicación biológica sobre los procesos evolutivos que moldearon a las poblaciones en estudio, aunque se utilizan en estudios de diversidad para fines de comparación (De Vicente et al., 2003).

2.4.2.3. Visualización de las relaciones

Para la visualización de las relaciones se necesitan métodos que realicen mediciones de más de una variable a la vez por cada individuo o unidad de muestreo, a estos métodos se les denomina métodos multivariados. Los métodos multivariados se dividen en dos grandes clases: La clasificación (o agrupamiento) y la ordenación.

La clasificación se refiere a un grupo de técnicas multivariadas cuyo objetivo principal es el agrupamiento de individuos o unidades de muestreo basado en las características que ellas poseen, de manera que los individuos más similares se agruparán matemáticamente en el mismo grupo (Mohammadi y Prasanna, 2003). Los métodos de agrupamiento se dividen a su vez en dos categorías:

- (i) Los métodos basados en distancia, en donde el algoritmo de agrupamiento necesita primero de una matriz de distancia por pares para poder dar como resultado una representación gráfica conocida como árbol o dendrograma.
- (ii) Los métodos basados en modelos, en donde se asume que las observaciones de cada grupo son reflejos de un modelo paramétrico, y las inferencias sobre los parámetros correspondientes a cada grupo y la membresía de cada individuo a determinado grupo son realizadas usando

métodos estadísticos estándares como la máxima verosimilitud y el método bayesiano, por ejemplo, un método de agrupamiento basado en modelos utilizando estadística bayesiana fue elaborado por Pritchard y colaboradores (2000), el cual se encuentra implementado en el software Structure, puede realizar gráficos tipo histograma en donde cada grupo es representado por un color y los individuos o unidades de muestreo pueden tener un color o más (*admixture model*).

Los métodos basados en distancia pueden ser categorizados en dos grupos: jerárquicos y no jerárquicos, aunque en la literatura también se utiliza el término particional en vez de no jerárquico (Zhong y Ghosh, 2003). Dentro de los métodos de agrupamiento jerárquicos se encuentran los métodos denominados jerárquicos aglomerativos en donde se inicia con un número de grupos igual al número de individuos los cuales son agrupados de acuerdo a su similitud hasta agruparlos en un cierto número de grupos, por ejemplo, el método de agrupamiento de grupos de pares no ponderados usando la media aritmética o UPGMA, el cual minimiza las distancias entre los grupos tomando el promedio de la distancia por pares entre todos los individuos de la muestra (De Vicente et al., 2003). El método de agrupamiento por K-medias es un ejemplo clásico de método de agrupamiento no jerárquico o particional (Zhong y Ghosh, 2003), en donde se dividen los datos en K grupos que es especificado a priori de acuerdo a un criterio de optimización.

La ordenación, a diferencia de la clasificación, ordena a las muestras a lo largo de una gradiente en vez de dividir las en grupos. Estos métodos se utilizan bastante en ecología de comunidades debido a la naturaleza continua de los ecosistemas (Palmer, 2004) y porque estos métodos representan enfoques cercanos a la realidad biológica (De Vicente et al., 2003). Los métodos de ordenación reducen la situación real a un espacio dimensional bajo para que pueda ser visualizado en dos o tres dimensiones ya que es imposible visualizar múltiples dimensiones (4 o más) simultáneamente en el espacio (Palmer, 2004).

Los métodos de ordinación pueden estar basados en datos de distancia o basados en eigenvalores (la suma de todas las varianzas para cada carácter en cada componente), aunque esta distinción es arbitraria debido a que algunos métodos de ordinación basados en distancia pueden ser resueltos a través de eigenvalores y técnicas basadas en eigenvalores pueden ser descritas en un marco de trabajo basado en distancia (Palmer, 2004).

El análisis de componentes principales (PCA) y el análisis de coordenadas principales (PCoA), los métodos de ordinación más conocidos, son utilizados para derivar un gráfico de 2 o 3 dimensiones donde se representan a los individuos en el espacio de tal manera que las distancias geométricas entre los individuos en el gráfico reflejan las distancias genéticas entre estos con una mínima distorsión, sin embargo, la diferencia entre estos dos métodos de ordinación es que el PCoA utiliza inicialmente una matriz de disimilitud o similitud mientras que el PCA utiliza la matriz de datos iniciales como los datos de presencia o ausencia de alelos de los marcadores moleculares (Mohammadi y Prasanna, 2003).

El PCA reduce los datos por transformación lineal de las variables originales en un nuevo conjunto de variables no correlacionadas denominadas componentes principales que resumen la variabilidad presente en los datos originales, siendo el primer componente principal el que mejor resume la variabilidad seguido del segundo componente, seguido del tercero y así sucesivamente. Es preferible que los primeros componentes puedan explicar de manera correcta la mayoría de la variabilidad en los datos para que los primeros componentes (2 o 3) puedan ser usados como ejes en el gráfico (Mohammadi y Prasanna, 2003). Otros métodos de ordinación incluyen la ordenación polar, escalado multidimensional no métrico (NMDS), análisis de correspondencia, etc (Palmer, 2004).

3. HIPÓTESIS

Las 129 accesiones de *Chenopodium quinoa* Willd. de la región Puno presentan altos índices de diversidad genética.

4. OBJETIVOS

4.1. Objetivo general

Estudiar la diversidad genética de 129 accesiones de *Chenopodium quinoa* Willd. de la región Puno utilizando 15 marcadores moleculares microsatélites.

4.2. Objetivos específicos

- Cuantificar la diversidad genética de las 129 accesiones mediante los índices de diversidad.
- Identificar una posible estructuración poblacional en las 129 accesiones y determinar el grado de diferenciación genética entre los grupos.
- Medir la relación genética entre las 129 accesiones mediante la distancia de Bruvo para visualizarlo en un gráfico de dispersión.
- Identificar accesiones duplicadas entre las 129 accesiones.

5. MATERIALES Y MÉTODOS

5.1. Material vegetal

El material vegetal utilizado forma parte de la colección nacional de quinua que se encuentra conservado en el banco de germoplasma de la estación experimental agraria Illpa en Puno del Instituto Nacional de Innovación Agraria (INIA). Las 129 accesiones utilizadas en esta investigación provienen de 5 de las 13 provincias del departamento de Puno (San Román, Puno, El Collao, Chucuito y Yunguyo), y 13 distritos (Cabana, Puno, Paucarcolla, Vilque, Tiquillaca, Pichacani, Acora, Ilave, Pisacoma, Huacullani, Juli, Pomata y Yunguyo) de estas 5 provincias, con 22 localidades geográficas diferentes (Figura 5.1). Estas accesiones presentaban mayor información de caracterización agronómica. Las accesiones se encuentran codificadas con un código nacional y un código de accesión (Tabla 5.1).

Se sembraron 8 semillas por accesión en bandejas de siembra de color negro (en cada compartimiento iba un tipo de accesión diferente) utilizando sustrato Premix 5, en uno de los invernaderos de la sede central del INIA ubicado en el distrito de La Molina en Lima, Perú. Se cosecharon plántulas de 2 semanas después de la germinación y se almacenaron en bolsas de polietileno de 10 x 13 cm. Estas bolsas se almacenaron en el ultracongelador a -80 °C hasta su procesamiento para la extracción del ADN genómico. Todos los procedimientos subsiguientes se realizaron en el laboratorio de biología molecular y genómica ubicado también en la sede central del INIA (Figura 5.2).

5.2. Extracción del ADN genómico

Se utilizó las hojas y tallo de las plántulas de quinua para extraer su ADN genómico. El protocolo de extracción se basó en el utilizado por Doyle (1991) con ligeras modificaciones empleadas por Via y Rada (2015):

- Se agregó 1 ml del tampón de extracción CTAB 2X (Tabla 5.2) y 2.85 μ l de β -mercaptoetanol (se combina con el tampón inmediatamente antes de su uso para disminuir la posibilidad de oxidación) a cada bolsa con 100 mg de material

vegetal congelado y se trituró de forma mecánica utilizando un tubo de ensayo grueso.

- Se colectó la mezcla triturada de cada accesión en dos microtubos de 1.5 ml para así tener un respaldo de ADN genómico por cada accesión (2 microtubos con ADN genómico extraído por accesión al finalizar la extracción).
- Cada microtubo fue incubado a 65 °C por 45 minutos en constante agitación utilizando la incubadora Thermomixer® de Eppendorf.
- Posterior a la agitación, se añadió 900 ul de la solución cloroformo – alcohol isoamílico en la proporción 24:1 en cada microtubo, se homogenizó la mezcla en un agitador de vórtice Vortex-Genie 2 de Scientific Industries Inc. y se centrifugó a 14000 rpm durante 5 minutos utilizando una microcentrífuga refrigerada Mikro 22R de Hettich® a 4 °C.
- Se transfirió el sobrenadante a un nuevo microtubo de 2 ml y se adicionó 60 ul de tampón de extracción CTAB 10X entibiado (Tabla 5.2), se invirtió suavemente y se incubó a 65 °C por 5 minutos.
- Se agregó nuevamente 900 ul de cloroformo – alcohol isoamílico, se homogeneizó y se centrifugó a 14000 rpm por 5 minutos. Se recuperó el sobrenadante y se colocó en nuevos microtubos de 1.5 ml.
- Se adicionó isopropanol frío (-20 °C) en un volumen igual al del sobrenadante y se colocaron las muestras a -20 °C por 30 minutos.
- Después se centrifugaron los microtubos a 14000 rpm por 20 minutos, se decantó los sobrenadantes y se invirtieron los microtubos sobre papel toalla durante 1 minuto.
- Se realizó un lavado con 1 ml de etanol 70% frío por microtubo y se centrifugó a 14000 rpm por 5 minutos. Se decantó el sobrenadante y se dejó secar para eliminar el remanente del etanol y obtener un precipitado completamente seco.
- Se añadió 300 ul de una solución de acetato de potasio 3 M frío y se incubó a -20 °C por 10 minutos.

- Luego se centrifugó a 14000 rpm por 5 minutos, se eliminó el sobrenadante y se realizó otro lavado con etanol, pero de diferente concentración (etanol 90%), se centrifugó a 14000 rpm por 5 minutos y se decantó el sobrenadante.
- Se invirtieron los microtubos en papel toalla y se esperó hasta que el etanol se evapore completamente.
- Obtenido un precipitado de ácidos nucleicos completamente seco, se adiciona 100 ul de agua libre de nucleasas y se deja disolver a 4 °C.
- Si la calidad del ADN era buena (determinado por electroforesis), entonces se añadía 2 ul de ARNasa 10 mg/ml y se incubaba a 37°C por 1 hora. Los microtubos se almacenaron a -20 °C para su conservación y posterior uso.

La calidad del ADN genómico fue comprobada por electroforesis en gel de agarosa al 1%. Se colocaron 2 ul de la solución madre de ADN genómico (microtubos con 100 ul de ADN genómico producto de la extracción) más 8 ul de tampón de carga SALB 1X (Tabla 5.2) en cada pocillo del gel para la corrida electroforética a voltaje constante de 100 V por 20 minutos utilizando un tampón de corrida TBE 1%. En la preparación de los geles se agregaba bromuro de etidio 10 mg/ml como agente intercalante del ADN para revelar las bandas de ADN utilizando luz ultravioleta. Para observar las bandas en el gel se utilizó el fotodocumentador ChemidocTM XRS+ de Bio-rad (Figura 5.3a).

5.3. Cuantificación del ADN genómico

El ADN genómico de las 129 accesiones fue cuantificado utilizando el espectrofotómetro de microplacas Epoch de Biotek®. Se utilizó el programa Gen5TM para la obtención de valores como el índice de pureza 260/280 (la relación de las absorbancias a 260 y 280 nm) y las concentraciones de ADN en ng/ul. Una vez seleccionadas las 129 soluciones madre de ADN genómico de mejor calidad, se siguió con la preparación de las soluciones de trabajo de ADN genómico (20 ng/ul) utilizando agua libre de nucleasas para la dilución y placas de 96 pocillos profundos para el almacenamiento de las soluciones de trabajo a -20°C. Se comprobó la calidad de las soluciones de trabajo por

electroforesis en gel de agarosa al 2% (5 ul de solución de trabajo más 5 ul de tampón de carga SALB 2X, 120 V por 20 minutos) y se cuantificó la concentración de éstas utilizando el espectrofotómetro.

5.4. Selección de marcadores microsatélites y amplificación por PCR

Se seleccionaron 15 marcadores microsatélites (Tabla 5.3) de los 23 marcadores utilizados por Vía y Rada (2015) según su riqueza alélica (número de alelos), heterocigosidad esperada y observada. De estos 15 marcadores, 11 (excepto KAAT006, QAAT76, QAAT70, QGA032) se encuentran en distintos grupos de ligamiento del primer mapa genético de *C. quinoa* basado en microsatélites, lo que asegura un análisis más robusto a pesar del alto número de grupos de ligamientos (38 grupos que cubren 913 cM segun Jarvis et al., 2008). Las reacciones de amplificación se llevaron a cabo en placas de PCR de 96 pocillos, se ajustaron a un volumen final de 10 ul y consistieron de 40 ng de ADN genómico, 1X de solución de tampón de PCR de Kapa Biosystems, 2.25 mM de MgCl₂, 0.2 mM de dNTPs, 0.4 mg/ml de BSA grado molecular (albúmina sérica bovina), 0.35 U de Taq polimerasa *hot-start* de Kapa Biosystems, 0.06 – 0.1 uM del cebador universal M13F (-29) (5'-CACGACGTTGTAAAACGAC-3') unido a uno de los 4 fluoróforos utilizados en el sistema de electroforesis capilar ABI 3130xl (6-FAM, VIC®, NED®, PET®) y 0.02 – 0.1 uM de cada cebador directo y reverso (Tabla 5.4 y 5.5). La concentración de los cebadores directos era menor con respecto a los cebadores reversos con el fin de que se agoten y así los cebadores universales tomen su lugar, incorporando los fluoróforos a los amplificados (Culley et al., 2013). Todos los cebadores directos tenían la secuencia del cebador universal en el extremo 5' según el método descrito por Schuelke (Schuelke, 2000). Los pares de marcadores QGA021/QGA024, QAAT027/QAAT112 y QAAT74/QAAT100 se amplificaron en una misma reacción como PCR multiplex (Tabla 5.5), mientras que las demás reacciones tuvieron solo un marcador por reacción. Las amplificaciones se realizaron en el termociclador Mastercycler® pro de Eppendorf, siguiendo el siguiente protocolo: desnaturalización

inicial a 94 °C por 3 minutos; 30 ciclos de 94 °C por 1 minuto, una temperatura de hibridación específica para un par de cebadores por 1 minuto y 72 °C por 1 minuto; seguido de 8 ciclos de 94 °C por 30 segundos, 53 °C (temperatura de hibridación para el cebador universal) por 30 segundos y 72 °C por 30 segundos; y una extensión final a 72 °C por 10 minutos. Las amplificaciones fueron comprobadas por electroforesis en gel de agarosa al 2% (2 ul del amplificado más 2 ul de tampón de carga Laemmli 2X, 120 V por 30 minutos).

5.5. Separación de productos amplificados por electroforesis capilar

Los productos de PCR fueron separados en el analizador genético ABI 3130xl de Applied Biosystems™ (Figura 5.3b) mediante electroforesis capilar utilizando el gel polímero POP-7 a ciertas condiciones (Tabla 5.6). Para el sistema de electroforesis capilar ABI 3130xl, existen 4 canales diferentes para los 4 fluoróforos utilizados (6-FAM, VIC®, NED® y PET®) sin contar el canal para el fluoróforo unido el marcador estándar de tamaño (LIZ), además, este sistema permite una inyección simultánea de 16 muestras diferentes utilizando 16 capilares y la automatización de 6 corridas electroforéticas máximo. Cada producto amplificado con un tipo diferente de fluoróforo unido (Tabla 5.7) fue mezclado con el agente desnaturalizante formamida altamente desionizada (Hi-Di) y un marcador interno alélico 600 LIZ® de 36 fragmentos de rango 20 a 600 pb. Los volúmenes de cada componente eran los siguientes: 0.8 ul de producto amplificado unido a un tipo de fluoróforo, 8.6 ul de formamida y 0.4 ul de 600 LIZ®, resultando en un volumen final máximo de 12.2 ul por pocillo. Para la mezcla de los productos amplificados, formamida y marcador estándar de tamaño se utilizaron placas para el analizador genético de 96 pocillos. Una vez hecha la mezcla, la placa era sometida a una incubación a 95 °C por 3 minutos seguido de una incubación a -20 °C por 5 minutos en la congeladora para la obtención de fragmentos de ADN de una sola cadena. En caso de tener productos amplificados que se solapaban en tamaño en un mismo pocillo, éstos tenían que estar unidos a diferentes fluoróforos para su diferenciación en los

electroferogramas (Guichoux et al., 2011). Los resultados de la electroforesis fueron almacenados en archivos de formato FSA (*fragment analysis data*), visualizados y editados en el programa GeneMapper® versión 4.0, de tal forma que a cada pico de RFU se le asignó su respectivo tamaño en pb por comparación con el marcador estándar, obteniéndose electroferogramas con el tamaño en pb en el eje X y la intensidad de la fluorescencia (en RFU) en el eje Y.

5.6. Genotipado de los productos amplificados

La asignación de tamaño a cada pico, la discriminación entre picos o ruido, la discriminación entre picos verdaderos o artefactos y la determinación de los fenotípicos alélicos se realizaron en el programa GeneMapper®. Se cambió el valor del límite de detección (LOD) de 50 RFU a 100 RFU para que todo pico debajo de este valor no fuera considerado como posible alelo (Butler, 2015). Asimismo, el valor por defecto de la razón de tartamudeo (*stutter ratio*) fue cambiado de 0.9 a 0.2 de acuerdo con lo recomendado (Butler, 2015; Leclair et al., 2004) para que los picos tartamudos cuya relación con los picos verdaderos fuera menor que 0.2 no sean considerados verdaderos alelos. El software utilizado asignó de manera automatizada la categoría de alelo (*allele calling*) a cada pico en el electroferograma que lograba pasar todos los umbrales mínimos (LOD, *stutter ratio*, PHR, etc.), sin embargo, se revisó todos los electroferogramas para verificar si los alelos habían sido asignados de manera correcta o si algunos picos no eran categorizados como alelos, y así poder corregir cualquier error de forma manual. Sólo se consideraron los picos cuyo tamaño en pb se encontraba en el rango citado de la bibliografía (Vía y Rada, 2015). Los diferentes alelos fueron codificados de acuerdo a su tamaño en pb y los datos perdidos fueron codificados como -9. Una vez corregido los errores se exportó la matriz de alelos a una tabla de Excel que sirvió de punto de partida para su uso en los programas bioinformáticos.

5.7. Análisis de datos

La matriz de alelos de Excel fue convertida en dos documentos de texto con diferente formato para su conversión a objetos *genambig* y *genind* para los paquetes Polysat (Clark y Jasieniuk, 2011) y Adegenet (Jombart, 2008) respectivamente, del software R (R Development Core Team, 2008), utilizando RStudio versión 1.0.136. El objeto *genambig* puede guardar información de ploidía y sirvió de base para su exportación como documento de texto para su utilización en programas como Structure (Clark, 2016a).

El equilibrio de Hardy-Weinberg y las heterocigosidades observadas no se determinaron debido a que no se contaba con la información completa de la dosis alélica, una de las dificultades que se presenta cuando se trabaja con organismos poliploides (Dufresne et al., 2014).

5.7.1. Determinación de grupos genéticos

Para determinar la existencia de una estructura poblacional que describa a las 129 accesiones de quinua de la región Puno se utilizó un método de agrupamiento basado en un modelo bayesiano y un análisis discriminante de componentes principales (DAPC) (Jombart et al., 2010) utilizando el programa Structure versión 2.3.4 (Pritchard et al., 2000) y el paquete Adegenet respectivamente.

En el caso de Structure, se empleó un modelo mixto de frecuencias alélicas correlacionadas con 20 repeticiones máximo por cada número de grupos permitidos en el análisis (K), utilizando información de las 22 localidades con la función LOCPRIOR (Hubisz et al., 2009), 10000 iteraciones burn-in y 50000 pasos de cadenas de Markov Monte Carlo (MCMC). Se evaluó el parámetro r para saber si la información de las localidades fue informativa (Bolton et al., 2016). Se realizaron 2 evaluaciones distintas variando el rango de valores K (1-10, 1-7) para tener un indicio en la primera evaluación de cuál era el valor óptimo de K. Este valor óptimo fue determinado observando los gráficos de la probabilidad logarítmica de los datos ($\ln P(K)$) y de los valores ΔK (Evanno

et al., 2005) utilizando Structure Harvester Web versión 0.6.94 (Earl y vonHoldt, 2012). Si hubiera diferentes valores de K en las diferentes evaluaciones, se prefiere el menor valor que capture la mayor estructura en los datos (Pritchard et al., 2000).

Como un método alternativo al realizado por Structure, se realizó el DAPC que es un método de ordinación libre de modelos genéticos. La ventaja reside en que no depende de ningún modelo de genética de poblaciones en particular para hallar el número de grupos genéticos (Bolton et al., 2016). El DAPC provee una eficiente descripción de grupos genéticos utilizando unas pocas variables sintéticas (llamadas funciones discriminantes) y minimiza la variación dentro de los grupos para maximizar la variación entre los grupos (Roullier et al., 2013).

Se utilizó la función *find.clusters* (del paquete Adegenet) que convierte los datos alélicos en componentes principales no correlacionados y utiliza un algoritmo de agrupamiento de k-medias para comparar las diferentes soluciones de agrupamiento usando el criterio de información bayesiano (BIC). El menor valor de BIC indica el número óptimo de grupos (Jombart y Collins, 2016). Se retuvo todos los componentes principales para el análisis con *find.clusters*.

Seguidamente, se realizó el DAPC utilizando la función *dapc* (del paquete Adegenet), la cual toma los grupos encontrados por *find.clusters* y maximiza la distancia entre estos. El DAPC depende de la transformación de los datos utilizando PCA como un paso previo, sin embargo, retener demasiados componentes principales puede llevar a un ajuste excesivo de las funciones discriminantes, lo que podría modelar cualquier estructura y discriminar virtualmente cualquier conjunto de grupos. Por esta razón, se calculó el valor α (α -score) que mide la diferencia entre la proporción de reasignaciones exitosas del análisis (discriminación observada) y valores obtenidos usando grupos aleatorios (discriminación al azar). La función *optim.a.score* (del paquete Adegenet) se utilizó para elegir el número óptimo de componentes principales sobre la base del valor de α -score más alto. Se corrió 100 simulaciones por aumento de 40 componentes principales.

Se consideró una asignación correcta de una accesión a un grupo si la probabilidad de membresía asociada era mayor de 0.8, mientras que los individuos con probabilidades de membresía a un grupo menor de 0.8 fueron considerados mixtos (Roullier et al., 2013).

5.7.2. Asignación de alelos a isoloci y determinación del tipo de herencia

Como la quinua es un organismo alotetraploide, se utilizó una propiedad matemática de los datos de marcadores codominantes cuando son registrados como variables alélicas binarias (presencia/ausencia): dos alelos estarán correlacionados negativamente si pertenecen al mismo locus y no correlacionados si pertenecen a diferentes loci cuando la población es infinita y se encuentra bajo apareamiento aleatorio (Clark y Schreier, 2015). Este enfoque se puede utilizar para asignar alelos a sus respectivos isoloci en organismos aloploidos o autoploidos diploidizados. Antes de llevar a cabo la asignación de alelos a isoloci se corroboró la existencia de accesiones duplicadas (mismo genotipo) ya que interfiere con la utilización del método de asignación, para esto se utilizó la función *assignClones* (del paquete Polysat), que hace uso de una matriz de distancia entre individuos para agruparlos en grupos de individuos asexualmente relacionados. Se utilizó la distancia de Bruvo por defecto mediante la función *meandistance.matrix* (del paquete Polysat) y un umbral (argumento *threshold*) de valor 0 para identificar accesiones idénticas (Clark, 2016b). También se dividió a las 129 accesiones según los grupos obtenidos por el DAPC ya que las poblaciones genéticamente estructuradas son poco apropiadas para el método de asignación (Clark y Schreier, 2015).

Para la asignación de alelos a isoloci, se utilizó la función *processDatasetAllo* (del paquete Polysat) que utiliza a su vez dos funciones: *alleleCorrelations*, que busca correlaciones negativas entre los alelos y utiliza estas correlaciones para realizar asignaciones preliminares; y *testAlGroups*, que ajusta estas asignaciones si es necesario después de compararlas con los genotipos individuales (Clark, 2016b). Se

utilizaron tres grupos de parámetros aparte de los cuatro grupos de parámetros por defecto que presenta la función *testAllGroups* para ajustar algunos parámetros a nuestros datos (Clark y Schreier, 2015; Clark, 2016b). Luego, con los mejores grupos de asignaciones alélicas, se recodificaron los datos y se evaluó la cantidad de genotipos perdidos por las asignaciones. Los resultados gráficos de la función *processDatasetAllo* y la cantidad de genotipos perdidos sirvieron de base para establecer los siguientes métodos que se utilizarían como el coeficiente de distancia y el método de estimación de frecuencias alélicas.

5.7.3. Estimación de frecuencias alélicas e índices de diversidad genética

Para la estimación de las frecuencias alélicas, se utilizó el método descrito por De Silva para la estimación de frecuencias alélicas bajo herencia polisómica y una tasa de autogamia conocida (De Silva et al., 2005). La tasa de autogamia utilizada fue de 0.9 (Bonifacio et al., 2015) y se utilizó la función *deSilvaFreq* del paquete Polysat. A partir de las frecuencias alélicas estimadas, se determinó la heterocigosidad esperada (H_s) de cada grupo genético encontrado por el DAPC y la H_s total de las 129 accesiones de quinua convirtiendo las frecuencias alélicas en un objeto *genpop* y utilizando la función *Hs* de Adegenet para hallar las heterocigosidades esperadas (Jombart, 2008). La heterocigosidad esperada por locus (H_{exp}) se estimó utilizando la función *summary* del paquete Adegenet. El número de alelos por locus y por grupo genético y el número promedio de alelos por locus (A) se estimaron utilizando la función *alleleDiversity* del paquete Polysat. El número de alelos privados por locus y por grupo se estimaron utilizando la función *private_alleles* del paquete Poppr versión 2.3.0 (Kamvar et al., 2014).

5.7.4. Distancia genética, índices de diferenciación genética y análisis de coordenadas principales

La matriz de distancia entre individuos se realizó utilizando la distancia de Bruvo (Bruvo et al., 2004) que asume un modelo de mutación por pasos y es independiente del nivel de ploidía de los organismos (Dufresne et al., 2014). El cálculo de la matriz se realizó utilizando la función *meandistance.matrix2* (paquete Polysat) que utiliza la función *genotypeProbs*, ésta calcula todos los posibles genotipos no ambiguos y sus probabilidades bajo apareamiento aleatorio o autogamia parcial, para esto utiliza las frecuencias alélicas estimadas por *deSilvaFreq*.

Asimismo, se estimó los índices de diferenciación poblacional F_{ST} (Wright, 1950) y G_{ST} (Nei y Chesser, 1983) entre los grupos genéticos determinados por el DAPC utilizando la matriz de distancia entre individuos y la función *calcPopDiff* del paquete Polysat.

El análisis de coordenadas principales (PCoA) se realizó utilizando la matriz de distancias y los grupos genéticos determinados por el DAPC.

6. RESULTADOS

6.1. Calidad del ADN genómico

La calidad del ADN genómico extraído fue buena de acuerdo a los índices de pureza 260/280 (Tabla 6.1). El índice promedio de pureza de las muestras de ADN genómico fue de 2.1 (SD = 0.08) en un rango de 1.89 – 2.31. La cantidad promedio de ADN genómico extraído fue de 218.997 ng/ul (SD = 85.58 ng/ul) en un rango de 67.98 – 586.83 ng/ul. La calidad también se corroboró mediante electroforesis en gel de agarosa donde se observaron bandas densas (Figura 6.1) lo que indica una buena extracción del ADN genómico.

6.2. Amplificación de las regiones microsatélites y separación de los productos amplificados

Todos los marcadores microsatélites amplificaron de acuerdo con las bandas obtenidas por la electroforesis en gel de agarosa (Figura 6.2). Los marcadores microsatélites dinucleótidos (QGA021, QGA024, QGA032 y KGA3) presentaron mayor cantidad de picos tartamudos que los demás marcadores trinucleótidos (Figura 6.3). Aunque el fluoróforo NED[®] presentó señales débiles con respecto a los otros fluoróforos, se pudo determinar los alelos para todos los marcadores unidos a NED[®]. El número de genotipos perdidos fue de 194, representando un 10.03% de todos los genotipos (1935 genotipos). El número mínimo y máximo de alelos encontrados por genotipo de cada marcador fue de 1 a 4, este rango de alelos concuerda con la alotetraploidía de la quinua.

6.3. Estructura poblacional

En el análisis con el programa Structure, se utilizó la metodología de Evanno y colaboradores (2005) para estimar el número de grupos apropiados. Se realizaron 2 evaluaciones: en la primera evaluación (20 repeticiones, K = 1-10) se encontró un valor máximo de ΔK (4.608) en K = 4 con un pico secundario en K = 6 (4.033) y en la segunda evaluación (20 repeticiones, K = 1-7, Tabla 6.2) se encontró un valor máximo de ΔK

(6.738) en $K = 3$ y un segundo valor máximo (5.024) en $K = 6$ (Figura 6.4). Se consideró el menor valor de K ($K = 3$) como el valor de K apropiado (segunda evaluación). El valor promedio del parámetro r para $K = 3$ (segunda evaluación) fue de 11.3137 (SD = 3.4119) lo que indica que la información de las localidades no fue informativa según Hubisz et al. (2009).

El método de agrupamiento por k -medias implementado por la función *find.clusters* del paquete Adegenet encontró que el menor valor de BIC coincidió con el valor de $K = 3$ (Figura 6.5). El tamaño de los tres grupos hallados por k -medias, K_1 , K_2 y K_3 , fue de 42, 38 y 49 respectivamente (Figura 6.6a). La función *optim.a.score* halló el valor α -score óptimo (0.553) reteniendo 7 componentes principales (Figura 6.7). Entonces el DAPC se realizó reteniendo 7 componentes principales y todas las funciones discriminantes (2 eigenvalores). El tamaño de los tres grupos posteriores para los grupos K_1 , K_2 y K_3 fue de 39, 37 y 53 respectivamente (Figura 6.6b). La proporción total de asignaciones correctas fue de 0.969 y las proporciones de asignaciones correctas por grupo fueron de 0.9286, 0.9737 y 1 para los grupos K_1 , K_2 y K_3 respectivamente (Figura 6.6c). La proporción de la varianza conservada fue de 0.293, lo que se ve reflejado en el gráfico del DAPC en donde las accesiones de los 3 grupos no se encuentran tan alejadas en cuanto a distancia (Figura 6.8).

Se encontraron 21 accesiones que fueron consideradas accesiones mixtas de acuerdo a las probabilidades posteriores de membresía (Tabla 6.3). En cuanto a las accesiones no mixtas: Del grupo K_1 (Figura 6.9a), el 77% de las accesiones pertenecían a la provincia de Puno (distritos de Puno, Acora y Paucarcolla), el 14% pertenecían a la provincia de Chucuito (distritos de Pomata, Juli y Pisacoma) y el 9% restante pertenecían a las provincias de Yunguyo (distrito de Yunguyo) y San Román (distrito de Cabana); del grupo K_2 (Figura 6.9b), el 83% de las accesiones pertenecían a la provincia de Puno (distritos de Puno, Acora y Paucarcolla) y el 17% restante pertenecían a la provincia de Chucuito (distritos de Pomata, Huacullani y Pisacoma); y del grupo K_3 (Figura 6.9c), el 50% de las accesiones pertenecían a la provincia de Puno (distritos de

Puno, Acora, Paucarcolla, Pichacani, Tiquillaca y Vilque), el 39% pertenecían a la provincia de Chucuito (distritos de Pomata, Huacullani y Juli) y el 11% restantes pertenecían a las provincias de San Román (distrito de Cabana), El Collao (distrito de Ilave) y Yunguyo (distrito de Yunguyo).

Los gráficos de las probabilidades de membresía de las 129 accesiones generados por el DAPC y Structure para $K = 3$ (Figura 6.10) resultaron ser similares, aunque en el gráfico generado por Structure se puede observar mayor cantidad de individuos mixtos. Las probabilidades posteriores de membresía de las 129 accesiones de quinua a cada grupo (Tabla 6.3) generadas por el DAPC fueron utilizadas para dividir a la población en 3 grupos (K_1 , K_2 , K_3) y realizar las siguientes evaluaciones. Las accesiones mixtas no fueron descartadas para los siguientes análisis.

6.4. Determinación del tipo de herencia

Se dividió a las 129 accesiones en 3 grupos de acuerdo a los grupos encontrados por el DAPC para el uso de la función *processDatasetAllo*. Asimismo, se determinó la ausencia de clones dentro de la muestra, por lo que se pudo utilizar la función sin eliminar accesiones. La calidad del agrupamiento por k-medias de cada locus fue uniforme, aunque no se encontraron en el cuadrante superior derecho que indica un buen agrupamiento de los alelos en los isoloci (Figura 6.11). Ninguno de los gráficos de los valores P para las correlaciones entre los alelos de cada locus y grupo tuvo un patrón de división de los alelos en 2 isoloci (Figura 6.12) lo que se sugiere una herencia polisómica en vez de una herencia disómica estricta según Clark (2016b). Las mejores asignaciones de alelos a isoloci de cada locus tuvo una proporción promedio de alelos homoplásicos (alelos que se encuentran en ambos isoloci) de 33.22% (SD = 14.99 %), sin embargo, el marcador QAAT76 no tuvo una asignación adecuada por lo que no se pudo separar en dos isoloci. La recodificación de los datos aumentó la cantidad de genotipos perdidos (1937 genotipos perdidos de 3741 genotipos en total, 51.78%) a la mitad de los genotipos, por lo que no se optó por esta metodología. Basado en los

gráficos de valores P y la cantidad de genotipos perdidos generados, los datos fueron tratados asumiendo una herencia polisómica.

6.5. Índices de diversidad y diferenciación genética

Un total de 179 alelos fueron identificados por los 15 loci microsatélites entre las 129 accesiones de quinua analizadas (Tabla 6.4) y la heterocigosidad esperada (H_s) de las 129 accesiones utilizando las frecuencias alélicas estimadas por el método de De Silva fue de 0.7657. El número de alelos por locus (Tabla 6.4) varió de 7 (KAAT006) a 19 (QAAT100) con un promedio (A) de 11.93 alelos por locus. (Tabla 6.4). Los valores de heterocigosidad esperada (H_{exp}) por locus varió en un rango de 0.61 (KAAT006) a 0.87 (QGA024, QAAT76 y QAAT100) con un promedio de 0.77 (Tabla 6.4). De acuerdo a la definición de Ott (1992), los marcadores son considerados polimórficos si la $H \geq 0.1$ y altamente polimórficos si la $H \geq 0.7$, por lo que todos los marcadores utilizados fueron polimórficos y 11 de los 15 marcadores (73.33 %) fueron altamente polimórficos (Tabla 6.4).

Con respecto a los índices de diversidad para los grupos genéticos encontrados por el DAPC, los valores de heterocigosidad esperada (H_s) de cada grupo utilizando las frecuencias alélicas estimadas por el método de De Silva fue de 0.6957, 0.7198 y 0.7761 para los grupos K1, K2 y K3 respectivamente (Tabla 6.5). El grupo K3 presentó el mayor número de alelos (154 alelos) y un número promedio de alelos por locus de 10.27, mientras que el grupo K1 presentó el menor número de alelos (123 alelos) y un número promedio de alelos por locus de 8.2. El número de alelos del grupo K2 fue de 133 y el número promedio de alelos por locus fue de 8.87 (Tabla 6.5). Los números de alelos privados para los grupos K1, K2 y K3 fueron de 14, 15 y 48 respectivamente (Tabla 6.5). Los porcentajes de loci que presentaron alelos privados fueron de 40%, 60% y 86.67% para los grupos K1, K2 y K3 respectivamente.

Para examinar la diferenciación genética entre los 3 grupos genéticos encontrados, se calculó el índice de fijación F_{ST} y el coeficiente de diferenciación genética G_{ST} . Las

estimaciones de los valores por pares F_{ST} y G_{ST} se presentan en la tabla 6.6. Los valores de F_{ST} oscilaron entre 0.0288 y 0.0461. Los grupos K2 y K3 tuvieron el valor F_{ST} más cercano a 0 ($F_{ST} = 0.0288$), mientras que los grupos K1 y K3 tuvieron el valor más lejano ($F_{ST} = 0.0461$). Se encontraron resultados similares con el valor G_{ST} , que osciló entre 0.0263 y 0.0449.

6.6. Distancia de Bruvo y análisis de coordenadas principales

Las distancias genéticas entre individuos utilizando la distancia de Bruvo se utilizó para realizar un análisis de coordenadas principales (Figura 6.13). Los coeficientes de la distancia de Bruvo oscilaron entre los valores de 0.1741 y 0.8418 (valor promedio de 0.5473, SD = 0.0957), con la menor distancia entre las accesiones PER004527 y PER004528 (ambas accesiones miembros del grupo K2), mientras que la mayor distancia fue observada entre las accesiones PER004334 (del grupo K2) y PER004573 (del grupo K1). El PCoA de las 129 accesiones de quinua reveló un patrón de variación continuo entre las accesiones, en donde se observó una mayor agrupación entre las accesiones del grupo K3. Sin embargo, los tres grupos aparecen bien diferenciados uno del otro.

7. DISCUSIÓN

7.1. Calidad de la extracción del ADN genómico

Aunque los valores de índices de pureza de las extracciones de ADN genómico fueron más altos que el valor recomendado (1.8 para soluciones puras de ADN), valores más altos de pureza como los obtenidos en esta investigación no son indicativos de un problema como contaminación (Wilfinger et al., 1997).

Para la extracción del ADN genómico, se utilizaron 8 plántulas por cada accesión, mezclando el ADN genómico de estas plántulas que representaron a una sola accesión. A esta técnica se le denomina técnica del bulk (*leaf bulking strategy*), la cual es recomendable cuando los organismos no presentan una alta tasa de alogamia y por lo tanto no existe una variabilidad intra-accesión (Costa, 2014). Como la quinua presenta una tasa de alogamia que oscila de 8 a 10% (Christensen et al., 2007), se esperaría que la variabilidad intra-accesión sea baja. Esta variabilidad no supuso un problema en estudios anteriores (Christensen et al., 2007; Fuentes et al., 2009a; Fuentes et al., 2012) ya que se hallaron valores altos de diversidad, sin embargo, la investigación realizada por Costa (2014) comparó las dos técnicas (*bulk* y plantas individuales) y concluyó que existe una variabilidad intra-accesión que no es detectada por la técnica del *bulk*. Efectivamente, un estudio previo encabezado por Costa et al. (2012) arrojó altos índices de diversidad para una región pequeña y poco estudiada como lo es el noroeste de Argentina. Por otro lado, la técnica del *bulk*, aunque no detecte toda la variabilidad intra-accesión, sigue siendo una estrategia válida de trabajo que disminuye los costos en cuanto a reactivos y facilita la comparación de los resultados de este estudio con los anteriores realizados en quinua utilizando marcadores moleculares.

7.2. Dosis alélica y frecuencias alélicas

El número máximo de alelos encontrados por genotipo de todos los marcadores utilizados en este estudio fue de 4 alelos, lo cual concuerda con la naturaleza alotetraploide de la quinua (Kolano et al., 2016) y dos estudios previos en donde se

encontró amplificación de dos bandas distintas en varios de los marcadores desarrollados (Mason et al., 2005) y la aparición máxima de 4 picos por electroferograma (Via y Rada, 2015), confirmando la naturaleza alotetraploide (Ward, 2000b).

Además, anteriores investigaciones reportan que varios marcadores SSR amplificaron dos loci que segregaron de manera codominante (Maughan et al. 2004; Jarvis et al., 2008) y el mapa genético realizado por Jarvis y colaboradores reveló la presencia de marcadores SSR en dos loci y en diferentes grupos de ligamiento. Estos grupos de ligamiento representan cromosomas homeólogos putativos en el genoma alotetraploide de quinua (Jarvis et al., 2008). Sin embargo, Christensen y colaboradores indicaron que la heterogeneidad genética (presencia de múltiples alelos en un locus genético dentro de una sola accesión) era debido a la variedad intra-accesión que no es detectada por la técnica del bulk, técnica que ellos utilizaron empleando tres plántulas por accesión (Christensen et al., 2007). De igual manera, el trabajo de Fuentes y colaboradores (2009a) utilizaron 3 plántulas por accesión (técnicas del bulk) y se eligió trabajar con marcadores que amplificaban un solo amplicón (de los cuales 6 coinciden con los utilizados en el presente estudio: QAAT24, QAAT74, QAAT70, KGA3, QAAT50 y QAAT50), no obstante, los datos genotípicos fueron convertidos a una matriz binaria. Las matrices binarias son recomendadas cuando se utilizan datos de marcadores dominantes o datos codominantes en organismos poliploides, en donde existe una ambigüedad genotípica, por ejemplo, Veramendi y colaboradores (2013) también utilizaron una matriz binaria en la caracterización de la diversidad genética de la colección boliviana de quinua. Sin embargo, los datos tratados como dominantes reducen el contenido de información y limitan los análisis que toman en cuenta la heterocigosidad observada o distribución de frecuencias alélicas (Dufresne et al., 2014). Los dos únicos estudios donde se utilizaron plántulas individuales de quinua en contraste con la técnica de bulk (Costa et al., 2012; Costa, 2014) reportaron un patrón monogénico de bandas lo cual permitió trabajar los datos de manera codominante, sin embargo, la separación de los productos amplificados se realizó mediante una

electroforesis en gel de poliacrilamida al 6%, con una menor resolución que podría haber dificultado la detección de todos los alelos en comparación a la técnica de electroforesis capilar (Christensen et al., 2007; Fuentes et al., 2009a; Fuentes et al., 2012; Via y Rada, 2015).

En el presente estudio, el número máximo de alelos por electroferograma de todos los marcadores fue de 4, por lo que se asumió que todos los marcadores utilizados amplificaban dos loci por marcador. Un inconveniente a raíz del número máximo de 4 alelos por electroferograma es la presencia de ambigüedad genotípica dada por los heterocigotos parciales (2 ó 3 alelos por electroferograma en un individuo tetraploide), en donde no se sabe con certeza la dosis alélica de cada alelo identificado en un genotipo (Dufresne et al., 2014). Como no hay una determinación precisa de los genotipos, la utilización de algunas medidas de diversidad dependientes del genotipo puede quedar restringida, por ejemplo, la heterocigosidad observada, frecuencias alélicas y frecuencias genotípicas. Una solución a este problema ha sido la estimación de la dosis alélica basado en comparar las diferencias de intensidad en bandas de electroforesis o diferencias en las alturas de los picos de electroferogramas debido a una correlación de estas diferencias con el número de copias alélicas para cada alelo diferente (Esselink et al., 2004; Obbard et al., 2006). Sin embargo, para tener una estimación certera del número de copias de cada alelo se necesita de preparados de ADN genómico muy finos, lo cual es muy raro en estudios de ecología molecular o caracterización de germoplasma.

El problema de la incertidumbre de la dosis alélica se puede abordar utilizando un estimador de frecuencias alélicas basado en un enfoque de máxima verosimilitud y utilizando un algoritmo de Expectativa – Maximización (Dempster et al., 1977) desarrollada por De Silva y colaboradores (2005), en donde se estima las frecuencias alélicas utilizando los fenotipos alélicos presentes y tomando un conjunto de suposiciones como una herencia completamente disómica o polisómica, sin entrecruzamiento, un sistema de apareamiento mixto con una tasa de autogamia

conocida, presencia de alelos nulos, ninguna ventaja selectiva sobre un genotipo en particular y la población que se encuentra en equilibrio.

El programa SPAGeDi ofrece otro estimador de las frecuencias alélicas que, a diferencia del estimador de De Silva (2005), considera que cada alelo en un heterocigoto parcial tiene la misma probabilidad de estar presente en más de una copia (Hardy y Vekemans, 2002), representando una desventaja para este estimador ya que conlleva a una subestimación de las frecuencias de alelos comunes y sobreestimación de las frecuencias de alelos raros (Dufresne et al., 2014; Via y Rada, 2015). El paquete Polysat cuenta con una función que permite emplear el método de De Silva (Clark y Jasieniuk, 2011), sin embargo, su aplicación trabaja bajo la suposición de una herencia polisómica. Por tal motivo se tuvo que corroborar si la herencia de la quinua seguía un patrón disómico o polisómico para saber si se podía utilizar la función *deSilvaFreq* de Polysat. En la siguiente sección se explica cómo se dedujo, basado en los resultados de la función *processDatasetAllo* de Polysat, que la quinua presentaría una herencia polisómica o intermedia con tendencia a la polisomía, razón por la cual se optó por estimar las frecuencias alélicas utilizando el método de De Silva (2005).

7.3. Disómica vs. Polisómica

Como se había mencionado anteriormente, estudios basados en filogenética e hibridación genómica *in situ* ya han corroborado el origen alotetraploide de la quinua (Maughan et al., 2006; Kolano et al., 2011; Kolano et al., 2016). Sin embargo, el tipo de herencia no necesariamente está correlacionado con la naturaleza de un organismo poliploide. Un organismo de origen autopoliploide puede sufrir diploidización con el tiempo y presentar eventualmente una herencia más disómica que polisómica, del mismo modo, un organismo de origen alopoliploide con genomas progenitores cercanamente relacionados puede permitir la formación de multivalentes conllevando a una homogenización del genoma (Obbard et al., 2006; Meirmans y Van Tienderen, 2013). Una buena parte de la literatura (Bazile et al., 2015) indica que la quinua posee

una herencia disómica (Simmonds, 1971) para la mayoría de los rasgos cualitativos (Gandarillas, 1979b; Saravia, 1991; Silvestri y Gil, 2000). Sin embargo, Ward (2000b) describió tanto herencia disómica como tetrasómica en dos de los rasgos morfológicos analizando la segregación alélica en las generaciones F1 y F2; indicó que las proporciones de segregación tetrasómica observada en una minoría de familias pueden ser debido a un intercambio recíproco de fragmentos entre cromosomas homólogos.

La herencia intermedia, que hace referencia a una herencia entre la disómica y polisómica, es común en híbridos interespecíficos fértiles, dado que sus progenitores están muy relacionados y por lo tanto se espera que posean algún grado de homología cromosómica, por lo que en sistemas donde la hibridación es común se puede encontrar una herencia intermedia (Stift et al., 2008).

En quinua la hibridación es muy frecuente entre la forma domesticada y la maleza denominada *ajara* que crece alrededor de los cultivos en el altiplano (Bazile et al., 2013), y también se ha reportado la hibridación entre la quinua costera de Chile y otras especies como *Chenopodium hircinum* o *C. album* que crecen como maleza, dificultando a los agricultores para obtener líneas puras de cultivares de quinua costera (Fuentes et al., 2009a; Fuentes et al., 2012). Estos hechos sugieren que la herencia de la quinua podría ser intermedia.

Aunque no se tenga información sobre la tasa de tetrasomía (Stift et al., 2008) de la quinua que permita saber si la quinua se inclina más a una herencia disómica o polisómica, se ha encontrado que una baja tasa del intercambio alélico de un subgenoma a otro es suficiente como para homogeneizar las frecuencias alélicas en poliploides y, por lo tanto, los valores de diversidad genética y diferenciación genética entre poblaciones no se verían sesgados si se asumen que los datos poliploides se encuentran bajo herencia totalmente polisómica (Meirmans y Van Tienderen, 2013).

Si la tasa de tetrasomía conlleva directamente a un intercambio de alelos entre los subgenomas, entonces los valores de diversidad genética y diferenciación genética utilizando datos poliploides de quinua serían robustos. Los resultados que se obtuvieron

utilizando el algoritmo elaborado por Clark y Schreier (Clark y Schreier, 2015) para la asignación de alelos a isoloci en organismos aloploidos también sugiere que en la quinua existe una tasa moderada de intercambio de alelos entre isoloci. Si bien parte de los alelos compartidos se puede explicar por homoplasia de alelos, la cantidad de genotipos que no concuerdan con la asignación alélica es considerablemente alta como para que sea explicado solamente por homoplasia, lo que indicaría que la quinua presentaría una herencia polisómica o intermedia con tendencia a la polisomía.

Si bien éste podría representar la primera prueba de herencia intermedia de 15 SSR en *Chenopodium quinoa* utilizando el algoritmo de Clark y Schreier, sería conveniente y necesario realizar estudios de segregación (utilizando cruces, parentales y descendencia) de marcadores moleculares en quinua como ya se ha realizado en otras especies, por ejemplo, en especies de *Rorippa* (Stift et al., 2008) para determinar la tasa de tetrasomía de cada marcador microsatélite.

7.4. Estructura poblacional

La quinua ha sido clasificada en 5 ecotipos: quinua de los valles, de las yungas, del altiplano, de los salares y del nivel del mar (Tapia et al., 1980). Algunos estudios utilizando marcadores moleculares avalan esta división de ecotipos (Wilson, 1988; Costa, 2014; Vía y Rada, 2015), mientras que otros apuntan a una estructuración basada en la geografía (Veramendi et al., 2013). Sin embargo, la mayoría de estudios a una escala mayor revela la existencia de dos grupos representando dos grandes acervos genéticos: la quinua andina (*highlands*) y la quinua de la costa de Chile (*lowlands*) o kinwa (Wilson, 1998; Mason et al., 2005; Christensen et al., 2007; Del Castillo et al., 2007; Maughan et al., 2012; Fuentes et al., 2012; Costa, 2014).

En el presente estudio, todas las accesiones estudiadas provenían de la parte sur del departamento de Puno por lo que todas eran pertenecientes al ecotipo del altiplano (Risi y Galwey, 1984; Bazile et al., 2013) y las localidades geográficas no se utilizaron para definir subpoblaciones ya que no fueron informativas de acuerdo al valor del parámetro

r obtenido. Como no se tenía un criterio de división de las 129 accesiones en subpoblaciones, se tuvo que hallar los grupos genéticos utilizando el programa Structure y la función *dapc* del paquete R *adeigenet*, utilizando el método de agrupamiento basado en un modelo de agrupamiento bayesiano (Pritchard et al., 2000) y el método multivariado sin estar basado en un modelo (Jombart et al., 2010) respectivamente. Comparando ambos métodos, el método del DAPC arrojó mejores resultados debido a que no necesita realizar suposiciones como el equilibrio de Hardy-Weinberg o el equilibrio de ligamiento, además, es independiente del nivel del ploidía y la tasa de recombinación genética (Jombart et al., 2010).

El programa Structure trabaja sobre la base de suposiciones como el equilibrio Hardy-Weinberg y equilibrio de ligamiento por lo que los resultados utilizando Structure no fueron consistentes en cuanto al número de grupos. Cuando existen casos donde no hay un solo valor de K consistente, se sugiere que se tome el valor de K más bajo que capture la mayor estructura en los datos (Pritchard et al., 2010) que en el caso del presente estudio fue de 3.

El estudio de Vía y Rada (2015) también utilizó el método de agrupamiento k-medias (función *find.clusters* de *Adegenet*) para hallar grupos en las 172 accesiones de quinua del altiplano y valles interandinos y encontró 7 que no guardaban relación con la procedencia de la muestra por lo que se concluyó que el DAPC no arrojó una estructuración genética entre los grupos inferidos, sin embargo, los grupos 1 y 6 presentaron mayor cantidad de accesiones de los valles interandinos y estaban más próximos en el gráfico, los grupos 2, 4 y 7 que estaban constituidos principalmente por accesiones del altiplano también se encontraban en proximidad validando nuevamente la agrupación de la quinua por ecotipos.

También existe la posibilidad de que los grupos encontrados por el DAPC no sean discretos debido a una variación continua de la diversidad genética de la quinua del altiplano (Jombart et al., 2010). Christensen y colaboradores (2007) encontraron que el grupo de quinuas andinas (Ecuador, Perú, Bolivia y Argentina) no formaban subgrupos

discretos, pero si un continuo con los extremos superior e inferior conformados mayoritariamente por quinuas andinas del sur (Bolivia y Argentina) y del norte (Perú y Ecuador) respectivamente.

Los valores de diferenciación genética F_{ST} entre los tres grupos encontrados (min = 0.029, max = 0.046) también apoyan la hipótesis de una variación continua en los grupos. Los tres índices de fijación fueron menores de 0.05 (Tabla 6.6), lo que indica un nivel de flujo génico alto que no permitiría una diferenciación genética entre los grupos. Cabe resaltar que la explicación de los patrones de flujo génico entre las accesiones de quinua en el altiplano de Puno debe considerar las implicaciones generales de la intervención humana dada la asociación antigua de la quinua con el imperio incaico (Christensen et al., 2007).

La investigación llevada a cabo por Fuentes y colaboradores (2012) demuestra cómo el intercambio de semillas entre los agricultores y las prácticas locales de producción puede impactar en la estructura y diversidad genética a escala nacional en Chile. La variación continua de la diversidad en las 129 accesiones es más evidente en los gráficos de dispersión del DAPC (Fig. 6.8) y del PCoA (Fig. 6.13). En ambos gráficos no hay una formación de grupos totalmente discretos. Incluso el gráfico del PCoA, que utiliza la distancia de Bruvo que a su vez utilizó las frecuencias alélicas estimadas por el método de De Silva asumiendo equilibrio de Hardy-Weinberg (De Silva et al., 2005), no fue tan diferente del gráfico obtenido por el DAPC, que no asume equilibrio (Jombart et al., 2010), dándole soporte a la hipótesis de la continuidad de la variación genética de las accesiones de quinua del altiplano de Puno, además de validar la utilización del método de estimación de frecuencias alélicas de De Silva en el caso particular de la quinua.

7.5. Diversidad Genética

Se realizó la estimación de los índices de diversidad genética utilizando 4 medidas distintas: el número de alelos, número de alelos privados, la heterocigosidad esperada

de cada locus (H_{exp}) y la heterocigosidad esperada basado en las frecuencias alélicas obtenidas por el método de De Silva (H_s). Los valores de heterocigosidad esperada y número de alelos de cada locus de los 15 marcadores SSR utilizados en este estudio son comparados con los valores obtenidos en los 8 estudios previos de quinua utilizando marcadores SSR y recopilados en la tabla 7.1.

La diversidad alélica (número promedio de alelos por locus) de este estudio fue de 11.93 siendo similar (11.7) al obtenido por Christensen et al. (2007), fue mayor que el valor detectado para 59 accesiones del germoplasma de Chile (7.5) por Fuentes et al. (2009a) y menor que del valor de 35 accesiones del noroeste de Argentina (16) (Costa et al., 2012). El alto valor de diversidad alélica en la investigación de Costa et al. (2012) es debido a que se utilizaron 10 plantas individuales por accesión para evaluar la variación intra-accesión lo que permitió la detección de un mayor número de alelos. Via y Rada (2015) reportó un número promedio de alelos por locus de 12.78, mayor que el valor de este estudio debido al mayor número de accesiones y a que se analizó también quinuas de valles interandinos. De igual manera, el número de alelos encontrado en este estudio (179) es menor que el encontrado (294) por Via y Rada (2015).

El marcador con mayor número de alelos fue el locus QAAT100 ($A=19$), al igual que en el estudio de Via y Rada (2015) donde el mismo locus presentó el mayor número de alelos (44), mientras que el marcador con menor número de alelos fue KAAT006 con 7 alelos. El marcador QAAT100 también presentó el valor máximo de longitud (441 pb), lo que indicaría una alta tasa de mutación y explicaría el alto número de alelos encontrados (Goldstein y Clark, 1995), a pesar de lo reportado por Mason y colaboradores (2005), los cuales encontraron que no existe una correlación significativa entre la longitud del microsatélite y el número de alelos cuando el motivo es AAT.

De los 15 marcadores SSR utilizados, solo 6 obtuvieron valores de heterocigosidad mayores (QGA024, QAAT24, KGA3, QAAT76, QAAT74 y QAAT100) que las de las primeras publicaciones (Mason et al., 2005; Jarvis et al., 2008) y uno obtuvo el mismo valor (QAAT70, $H_{exp}=0.77$). A diferencia del valor H_{exp} , los valores del número de

alelos por locus en este estudio fueron mayores que los valores de las primeras publicaciones, a excepción del marcador QGA021 ($A=9$). Los estudios de Mason et al. (2005) y Jarvis et al. (2008) utilizaron paneles con accesiones de quinua de distintas partes de Sudamérica, abarcando una gran región geográfica y que representaba la diversidad de quinua en el continente, razón por la cual sus valores de heterocigosidad fueron mayores que en la de nuestro estudio.

Asimismo, los valores del número de alelos obtenidos en este estudio superaron los valores de las dos primeras investigaciones debido a que se utilizaron 129 accesiones a diferencia de las 31 y 23 accesiones utilizadas por Mason et al. (2005) y Jarvis et al. (2008) respectivamente.

Tanto el H_s de las 129 accesiones utilizando el método de De Silva (2005) como el Hexp promedio arrojaron un mismo valor de diversidad (0.77) y fue cercano al valor H promedio (0.75) de la investigación de Christensen y colaboradores (2007) quienes evaluaron 150 accesiones de 5 países sudamericanos, lo que indica un alto valor de diversidad genética y confirma la calidad altamente informativa de los marcadores SSR utilizados aquí (73% de los marcadores fueron altamente polimórficos). Paralelamente, este valor es menor al obtenido por Costa (2014), en donde se obtuvieron valores H_t (heterocigosidad total) de 0.89 y 0.82 para 45 accesiones de 5 países sudamericanos a excepción de Argentina y 35 accesiones del noroeste de Argentina respectivamente. Sin embargo, estos altos valores reportados en la tesis de Costa fueron debido a su metodología empleada (10 individuos por accesión). Por otro lado, la heterocigosidad esperada promedio en este estudio fue mayor que la de los estudios encabezados por Fuentes del año 2009 y 2012 en donde se analizaron 59 accesiones de Chile y 34 accesiones de diversos países respectivamente, cuyos valores fueron de 0.65 en ambos casos (Fuentes et al., 2009a; Fuentes et al., 2012).

Las quinuas del altiplano de Perú y Bolivia han presentado un valor de heterocigosidad mayor que las quinuas de otros ecotipos en estudios donde se ha utilizado accesiones de diversos países sudamericanos (Fuentes et al., 2012; Costa, 2014), lo que se traduce

en una alta diversidad genética de este ecotipo. No obstante, Via y Rada (2015) concluyó que 74 accesiones de quinua de los valles interandinos del Perú presentaban una mayor diversidad genética que 97 accesiones del altiplano del Perú basado en un mayor número de alelos privados en el grupo de los valles interandinos. Sin embargo, no se reportó el valor de heterocigosidad para los dos grupos (altiplano y valles interandinos). Si bien el número de alelos privados es otro indicador del nivel de diversidad genética, el valor de heterocigosidad podría corroborar si realmente la diversidad genética del ecotipo de valles interandinos supera al del ecotipo del Altiplano en Perú.

En cuanto a los índices de diversidad de los grupos obtenidos por el DAPC, el grupo K3 presentó la mayor diversidad genética según su valor de heterocigosidad y número de alelos privados (0.78, 48) mientras que los grupos K1 y K2 presentaron los menores índices de diversidad genética (Tabla 6.5). De igual manera, el grupo K3 también resultó ser el más diverso en cuanto a procedencia (Figura 6.9c) lo que se relaciona con la alta diversidad genética obtenida de este grupo. Cabe resaltar que la cantidad de alelos privados en el grupo K3 supera enormemente a los valores de los grupos K1 y K2 (14 y 15). Además, los valores de diferenciación genética (Tabla 6.6) indicaron que existe un mayor flujo génico entre los grupos K2 y K3 y un menor flujo génico entre los grupos K1 y K3.

Se considera que la zona del altiplano entre Puno y Bolivia representa el centro de origen de la quinua (Gandarillas, 1979b; Wilson, 1988) por lo que el nivel de diversidad genética es mayor en esta región (Christensen et al., 2007; Fuentes et al., 2009a; Costa, 2014). La alta diversidad genética de la quinua andina se refleja en el nivel de variación fenotípica que se encuentra en esta región de los Andes, alrededor del lago Titicaca entre Cuzco y el lago Poopó, presentando las quinuas diferentes colores de panoja, hojas, alturas de la planta, contenido de saponina, tipo de panoja y calidad del grano (Gandarillas, 1979a). Esta alta diversidad genética tiene relación con la cantidad de precipitaciones y la disponibilidad de agua en la zona. La investigación de Costa (2014)

encontró una relación negativa entre el nivel de precipitaciones y la diversidad genética. Como el altiplano es una región muy amplia, con escasez de lluvias y diversidad de hábitats debido a la compleja disponibilidad de agua en la zona, es de esperarse que la diversidad genética y fenotípica de la quinua andina sea muy alta.

8. CONCLUSIONES

- ❖ Las 129 accesiones de *C. quinoa* del banco de germoplasma del INIA procedente del departamento de Puno presentan una alta diversidad genética basado en los altos valores de heterocigosidad, número de alelos y alelos privados.
- ❖ Los análisis de estructuración poblacional no encontraron una estructura en la población estudiada.
- ❖ La herencia de los 15 marcadores SSR utilizados en las 129 accesiones es de carácter intermedia con tendencia a la polisomía.
- ❖ No se encontraron clones ni muestras duplicadas dentro de las 129 accesiones de quinua.
- ❖ Los marcadores SSR altamente polimórficos utilizados en este estudio son altamente eficientes para realizar estudios de diversidad genética en quinua.

9. RECOMENDACIONES

- ❖ Evaluar 10 plantas individuales por cada accesión de la colección nacional para conocer la variación intra-accesión.
- ❖ Utilizar otros cebadores universales aparte del utilizado en este trabajo (M13F(-29)) y en distintas combinaciones con los 4 fluoróforos (6-FAM, VIC®, NED®, PET®) para resolver el problema de la PCR con más de 2 cebadores (*multiplex*) donde los productos se solapan en longitud.
- ❖ Realizar estudios de segregación (utilizando cruces, parentales y descendencia) de marcadores SRR en quinua como ya se ha realizado en otras especies, por ejemplo, en especies de *Rorippa*, para determinar la tasa de tetrasomía de cada marcador microsatélite.
- ❖ Utilizar la información de la ausencia de clones para un futuro desarrollo de colecciones núcleo en el banco de germoplasma del INIA.
- ❖ Evaluar valores de heterocigosidad esperada de las accesiones de quinuas de los valles interandinos de la colección nacional para corroborar si realmente la diversidad genética del ecotipo de valles interandinos supera al del ecotipo del Altiplano en Perú.

10. REFERENCIAS BIBLIOGRÁFICAS

- ABUGOCH, L. E. Quinoa (*Chenopodium quinoa* Willd.): composition, chemistry, nutritional, and functional properties. *Advances in Food Nutrition Research*. 2009, vol 58, p. 1-31.
- ALLENDORF, F. W. y LUIKART, G. H. *Conservation and the Genetics of Populations*. 1era ed. Massachusetts, EE. UU: Blackwell Publishing, 2007, 642 p. ISBN-13: 978-1-4051-2145-3.
- ATWELL, W. A; PATRICK, B. M; JOHNSON, L. A. y GLASS, R. W. Characterization of quinoa starch. *Cereal Chemistry*. 1983, vol 60, num 1, p. 9-11.
- ANGIOSPERM PHYLOGENY GROUP (APG). An ordinal classification for the families of flowering plants. *Annals of the Missouri Botanical Garden*. 1998, vol 85, num 4, p. 531-553.
- BAZILE, D; FUENTES, F. F. y MUJICA, A. Historical perspectives and domestication. En: BHARGAVA, A. y SRIVASTAVA, S. *Quinoa: Botany, production and uses*. Boston, MA: CABI, 2013, p. 16-35. ISBN: 978-1-78064-226-0.
- BAZILE, D; BERTERO, D. y NIETO, C. *Estado del arte de la quinua en el mundo en 2013*. Montpellier, Francia: FAO (Santiago de Chile) y CIRAD, 2014, 724 p. ISBN: 978-92-5-308558-3.
- BAZILE, D; BERTERO, D. y NIETO, C. *State of the Art Report of Quinoa in the World in 2013*. Montpellier, Francia: FAO (Santiago de Chile) y CIRAD, 2015, 589 p. ISBN: 978-92-5-108558-5.
- BHARGAVA, A; SHULKA, S. y OHRI, D. Karyotypic studies on some cultivated and wild species of *Chenopodium* (Chenopodiaceae). *Genetic Resources and Crop Evolution*. 2006a, vol 53, num 7, p. 1309-1320.
- BHARGAVA, A; SHUKLA, S. y OHRI, D. *Chenopodium quinoa*: An Indian perspective. *Industrial Crops and Products*. 2006b, vol 23, num 1, p. 73-87.

- BHARGAVA, A; SHULKA, S. y OHRI, D. Genetic variability and interrelationship among various morphological and quality traits in quinoa (*Chenopodium quinoa* Willd.). *Field Crop Research*. 2007, vol 101, num 1, p. 104-116.
- BHARGAVA, A. y SRIVASTAVA, S. *Quinoa: Botany, production and uses*. Boston, Massachusetts: CABI, 2013, 247 p. ISBN: 978-1-78064-226-0.
- BIRD, C. E; KARL, S. A; SMOUSE, P. E. y TOONEN, R. J. Detecting and measuring genetic differentiation. En: HELD, C; KOENEMANN, S. y SCHUBART, C. (edit.). *Phylogeography and population genetics in Crustacea*. 1era ed. Boca Raton: CRC Press, Taylor and Francis Group, 2011, p. 31-55.
- BLACKET, M. J; ROBIN, C; GOOD, R. T; LEE, S. F. y MILLER, A. D. Universal primers for fluorescent labelling of PCR fragments – an efficient and cost-effective approach to genotyping by fluorescence. *Molecular Ecology Resources*. 2012, vol 12, num 3, p. 456-463.
- BOLTON, P. E; WEST, A. J; CARDILINI, A. P. A; CLARK, J. A; MAUTE, K. L; LEGGE, S; BRAZILL-BOAST, J; GRIFFITH, S. C. y ROLLINS, L. A. Three molecular markers show no evidence of population genetic structure in the Gouldian fish (*Erythrura gouldiae*). *Plos One*. 2016, vol 11, num 12, p. 1-19.
- BONIFACIO, A; PANDO-LUZ, L. y ROJAS, W. Quinoa breeding and modern variety development. En: BAZILE, D. et al. (edit). *State of the Art Report of Quinoa in the World in 2013*. Montpellier, Francia: FAO (Santiago de Chile) y CIRAD, 2015, p. 172-191. ISBN: 978-92-5-108558-5.
- BRUVO, R; MICHIELS, N. K; D'SOUZA, T. G. y SCHULENBURG, H. A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level. *Molecular Ecology*. 2004, vol 13, num 7, p. 2101-2106.
- BUTLER, J. M. *Forensic DNA Typing*. 2da ed. Massachusetts, EE. UU: Academic Press, 2005, 660 p. ISBN: 0-12-147952-8.
- BUTLER, J. M. *Advanced Topics in Forensic DNA Typing: Interpretation*. 1era ed. Oxford: Academic Press, 2015, 585 p. ISBN: 978-0-12-405213-0.

- CATACORA, A. G. Determinación del cariotipo en cinco líneas de quinua (*Chenopodium quinoa* Willd.). Tesis Título Profesional. Universidad Nacional Técnica del Altiplano, Puno, Perú, 1977.
- CHRISTENSEN, S. A; PRATT, D. B; PRATT, C; NELSON, P. T; STEVENS, M. R; JELLEN, E. N; COLEMAN, C. E; FAIRBANKS, D. J; BONIFACIO, A. y MAUGHAN, P. J. Assessment of genetic diversity in the USDA and CIP-FAO international nursery collections of quinoa (*Chenopodium quinoa* Willd.) using microsatellite markers. *Plant Genetic Resources*. 2007, vol 5, num 2, p. 82-95.
- CLARK, L. V. y JASIENIUK, M. Polysat: an R package for polyploid microsatellite analysis. *Molecular Ecology*. 2011, vol 11, num 3, p. 562-566.
- CLARK, L. V. y SCHREIER, A. D. Resolving microsatellite genotype ambiguity in populations of allopolyploid and diploidized autopolyploid organisms using negative correlations between alleles. *Molecular Ecology Resources*. 2015, p. 1-30.
- CLARK, L. V. *Polysat version 1.5 Tutorial Manual*. Illinois, EE. UU: Universidad de Illinois, 2016a, 54 p.
- CLARK, L. V. *Assigning alleles to isoloci in Polysat*. Illinois, EE. UU: Universidad de Illinois, 2016b, 27 p.
- COSTA, S. M; MANIFESTO, M. M; BRAMARDI, S. J. y BERTERO, H. D. Genetic structure in cultivated quinoa (*Chenopodium quinoa* Willd.), a reflection of landscape structure in Northwest Argentina. *Conservation Genetics*. 2012, vol 13, num 4, p. 1027-1038.
- COSTA, S. M. Variabilidad genética de *Chenopodium quinoa* Willd. en el noroeste argentino y su relación con la dispersión de la especie. Tesis Doctoral. Universidad Nacional de la Plata, Facultad de Ciencias Agrarias y Forestales, La Plata, Argentina, 2014.

- COULTER, L. y LORENZ, K. Quinoa-composition, nutritional value, food applications. *LWT-Food Science and Technology*. 1990, vol 23, p. 203-207.
- CULLEY, T. M; STAMPER, T. I; STOKES, R. L; BRZYSKI, J. R; HARDIMAN, N. A; KLOOSTER, M. R. y MERRITT, B. J. An efficient technique for primer development and application that integrates fluorescent labeling and multiplex PCR. *Applications in Plant Sciences*. 2013, vol 1, num 10, p. 1-10.
- DEL CASTILLO, C; WINKEL, T; MAHY, G. y BIZOUX, J. P. Genetic structure of quinoa (*Chenopodium quinoa* Willd.) from the Bolivian altiplano as revealed by RAPD markers. *Genetic Resources and Crop Evolution*. 2007, vol 54, num 4, p. 897–905.
- DEMPSTER, A. P; LAIRD, N. M. y RUBIN, D. B. Maximum likelihood from incomplete data via EM algorithm. *Journal of the Royal Statistical Society. Series B*. 1997, p. 1-38.
- DE SILVA, H. N; HALL, A. J; RIKKERINK, E; MCNEILAGE, M. A. y FRASER, L. G. Estimation of allele frequencies in polyploids under certain patterns of inheritance. *Heredity*. 2005, vol 95, num 4, p. 327-334.
- DE VICENTE, M. C; LOPEZ, C. y FULTON, T. *Genetic diversity analysis with molecular marker data: learning module*. IPGRI y Cornell, 2003.
- DOWNY, S. R; KATZ-DOWNIE, D. S. y CHO, K. J. Relationship of the Caryophyllales as suggested by phylogenetic analyses of partial chloroplast DNA ORF2280 homolog sequences. *American Journal of Botany*. 1997, vol 84, num 2, p. 253-273.
- DOYLE, J. J. DNA Protocols for Plants. En: HEWITT, G; JOHNSON, A. W. B. y YOUNG, J. P. W. (edit.). *Molecular techniques in taxonomy*. Berlin: Springer Berlin Heidelberg, 1991, p. 283-293. ISBN: 978-3-642-83962-7.
- DUFRESNE, F; STIFT, M; VERGILINO, R. y MABLE, B. K. Recent progress and challenges in population genetics of polyploid organisms: an overview of

- current state-of-the-art molecular and statistical tools. *Molecular Ecology*. 2014, vol 23, num 1, p. 40-69.
- EARL, D. A. y VONHOLDT, B. M. Structure Harvester: A website and program for visualizing Structure output and implementing the Evanno method. *Conservation Genetics Resources*. 2012, vol 4, num 2, p. 359-361.
- ESSELINK, G. D; NYBOM, H. y VOSMAN, B. Assignment of allelic configuration in polyploids using the MAC-PR (microsatellite DNA allele counting-peak ratios) method. *Theoretical and Applied Genetics*. 2004, vol 109, num 2, p. 402-408.
- EVANNO, G; REGNAUT, S. y GOUDET, J. Detecting the number of clusters of individuals using the software Structure: a simulation study. *Molecular Ecology*. 2005, vol 14, num 8, p. 2611-2620.
- EXCOFFIER, L; SMOUSE, P. E. y QUATTRO, J. M. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*. 1992, vol 131, num 2, p. 479-491.
- FAIRBANKS, D. J; BURGNER, K. W; ROBISON, L. R; ANDERSEN, W. R. y BALLON, E. Electrophoretic characterization of quinoa seed proteins. *Plant Breeding*. 1990, vol 104, num 3, p. 190-195.
- FAO. *Normas para bancos de germoplasma de recursos fitogenéticos para la alimentación y la agricultura*. Roma: Organización de las naciones unidas para la alimentación y la agricultura, 2014, 167 p. ISBN: 978-92-5-308262-9.
- FERREIRA, M. E. y GRATTAPAGLIA, D. *Introducción al uso de marcadores moleculares en el análisis genético*. 1era ed. Brasilia: EMBRAPA-CENARGEN, 1998, 200 p.

- FRANKHAM, R; BALLOU, J. D. y BRISCOE, D. A. *Introduction to Conservation Genetics*. 1era ed. New York, EE. UU: Cambridge University Press, 2002, 617 p. ISBN: 0-521-63985-9.
- FUENTES, F. F; MARTINEZ, E. A; HINRICHSEN, P. V; JELLEN, E. N. y MAUGHAN, P. J. Assessment of genetic diversity patterns in Chilean quinoa (*Chenopodium quinoa* Willd.) germplasm using multiplex fluorescent microsatellite markers. *Conservation Genetics*. 2009a, vol 10, num 2, p. 369-377.
- FUENTES, F. F; MAUGHAN, P. J. y JELLEN, E. N. Diversidad genética y recursos genéticos para el mejoramiento de la quinua (*Chenopodium quinoa* Willd.). *Revista geográfica de Valparaíso*. 2009b, num 42, p. 20-33.
- FUENTES, F. F. y BHARGAVA, A. Morphological analysis of quinoa germplasm grown under lowland desert conditions. *Journal of Agronomy and Crop Science*. 2011, vol 197, num 2, p. 124-134.
- FUENTES, F. F; BAZILE, D.; BHARGAVA, A. y MARTÍNEZ, E. A. Implications of farmers' seed exchanges for on-farm conservation of quinoa, as revealed by its genetic diversity in Chile. *The Journal of Agriculture Science*. 2012, vol 150, num 6, p. 702-716.
- GALLARDO, M; GONZÁLEZ, J. A. y PRADO, F. E. Presencia de betalaínas en plántulas de *Chenopodium quinoa* Willd. *Lilloa*. 2000, vol 40, num 1, p. 109-113.
- GALWEY, N. W. Exploited plants: quinoa. *Biologist*. 1989, vol 36, p. 267-274.
- GANDARILLAS, H. Botánica. En: TAPIA, M (edit). *Quinoa y Kañiwa: cultivos andinos*. Bogotá, Colombia: CIID, Oficina regional para la América Latina, 1979a, p. 20-44. ISBN: 0-88936-200-9.
- GANDARILLAS, H. Genética y origen. En: TAPIA, M (edit). *Quinoa y Kañiwa: cultivos andinos*. Bogotá, Colombia: CIID, Oficina regional para la América Latina, 1979b, p. 45-64. ISBN: 0-88936-200-9.

- GERBER, S; MARIETTE, S; STREIFF, R; BODÉNÈS C. y KREMER, A. Comparison of microsatellites and amplified fragment length polymorphism markers for parentage analysis. *Molecular Ecology*. 2000, vol 9, num 8, p. 1037-1048.
- GOLDSTEIN, D. B. y CLARK, A. G. Microsatellite variation in North American Populations of *Drosophila melanogaster*. *Nucleic Acids Research*. 1995, vol 23, num 19, p. 3882–3886.
- GONZÁLEZ, S. P. Estudio de flujo de genes en quinua (*Chenopodium quinoa* W.) en campo de agricultores mediante el uso de marcadores microsatélites. Tesis Título Profesional. Universidad de las Fuerzas Armadas (Escuela Politécnica del Ejército), Departamento de Ciencias de la Vida e Ingeniería en Biotecnología, Sangolquí, Pichincha, Ecuador, 2009.
- GUICHOUX, E; LAGACHE, L; WAGNER, S; CHAUMEIL, P; LÉGER, P; LEPAIS, O; LEPOITTEVIN, C; MALAUSA, T; REVARDEL, E; SALIN, F. y PETIT, R. J. Current trends in microsatellite genotyping. *Molecular Ecology Resources*. 2011, vol 11, num 4, p. 591-611.
- HARDY, O. J. y VEKEMANS, X. SPAGeDi: A versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Resources*. 2002, vol 2, num 4, p. 618-620.
- HUBISZ, M. J; FALUSH, D; STEPHENS, M. y PRITCHARD, J. K. Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*. 2009, vol 9, num 5, p. 1322-1332.
- HUGHES, A. R; INOUE, B. D; JOHNSON, M. T. J; UNDERWOOD, N. y VELLEND, M. Ecological consequences of genetic diversity. *Ecology Letters*. 2008, vol 11, num 6, p. 609-623.
- INSTITUTO NACIONAL DE INNOVACIÓN AGRARIA (INIA). Más de 50 líderes agrarios identifican variedades nativas de quinua resistentes a heladas [en línea]. 28 de setiembre del 2016, [5 de julio del 2017]. Disponible en Web:

<<http://www.inia.gob.pe/sala-de-prensa/notas-de-prensa/1775-2016-nota-n315>>.

INTEGRATED TAXONOMIC INFORMATION SYSTEM (ITIS). *Chenopodium quinoa* Willd. [en línea]. 2011, [ref. de 15 de setiembre del 2016]. Disponible en web: <https://www.itis.gov/servlet/SingleRpt/SingleRpt?search_topic=TSN&search_value=506567#>.

JARNE, P. y LAGODA, P. J. Microsatellites, from molecules to populations and back. *Trends in Ecology and Evolution*. 1996, vol 11, num 10, p. 424-429.

JARVIS, D. E; KOPP, O. R; JELLEN, E. N; MALLORY, M. A; PATTEE, J; BONIFACIO, A; COLEMAN, C. E; STEVENS, M. R; FAIRBANKS, D. J. y MAUGHAN, P. J. Simple sequence repeat marker development and genetic mapping in quinoa (*Chenopodium quinoa* Willd.). *Journal of Genetics*. 2008, vol 87, num 1, p. 39-51.

JELLEN, E. N; KOLANO, B. A; SEDERBERG, M. C; BONIFACIO, A. y MAUGHAN, P. J. *Chenopodium*. En: KOLE, C. (edit.). *Wild Crop Relatives: Genomic and breeding resources, legume crops and forages*. Berlín, Alemania: Springer, 2011, p. 35-61. eBook ISBN: 978-3-642-14228-4.

JIANG, G. L. Molecular markers and marked-assisted breeding in plants. En: Andersen, S. B. (edit.). *Plant breeding from laboratories to fields*. Intech, 22 de mayo del 2013, p. 45-83. ISBN: 978-953-51-1090-3.

JOHNSON, D. L. y WARD, S. M. Quinoa. En: JANICK, J. y SIMON, J. E (edit). *New Crops*. New York: Wiley, 1993, p. 219-221.

JOMBART, T. Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*. 2008, vol 24, num 11, p. 1403-1405.

JOMBART, T; DEVILLARD, S. y BALLOUX, F. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*. 2010, vol 11, num 1, p. 1-15.

- JOMBART, T. y COLLINS, C. *A tutorial for Discriminant Analysis of Principal Components (DAPC) using adegenet 2.0.2*. Londres, Inglaterra: Imperial College London, 2016, p. 1-43.
- KADEREIT, G; BORSCH, T; WELSING, K. y FREITAG, H. Phylogeny of Amaranthaceae and Chenopodiaceae and the evolution of C4 photosynthesis. *International Journal of Plant Science*. 2003, vol 164, num 6, p. 959-986.
- KALIA, R. K; RAI, M. K; KALIA, S; SINGH, R. y DHAWAN, A. K. Microsatellite markers: an overview of the recent progress in plants. *Euphytica*. 2011, vol 177, num 3, p. 309-334.
- KAMVAR, Z. N; TABIMA, J. F. y GRÜNWALD, N. J. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*. 2014, vol 2, p. 1-14.
- KOLANO, B; PANDO, L. G. y MALUSZYNSKA, J. Molecular cytogenetic studies in *Chenopodium quinoa* and *Amaranthus caudatus*. *Acta Societatis Botanicorum Poloniae*. 2001, vol 70, num 2, p. 85-90.
- KOLANO, B; GARDUNIA, B. W; MICHALSKA, M; BONIFACIO, D; FAIRBANKS, D; MAUGHAN, P. J; COLEMAN, C. E; STEVENS, M. R; JELLEN, E. N. y MALUSZYNSKA, J. Chromosomal localization of two novel repetitive sequences isolated from the *Chenopodium quinoa* Willd. *Genome Genome*. 2011, vol 54, num 9, p. 710-717.
- KOLANO, B; SIWINSKA, D; MCCANN, J. y WEISS-SCHNEEWEISS, H. The evolution of genome size and rDNA in diploid species of *Chenopodium s. l.* (Amaranthaceae). *Botanical Journal of the Linnean Society*. 2015, vol 179, num 2, p. 218-235.
- KOLANO, B; MCCANN, J; ORZECOWSKA, M; SIWINSKA, D; TEMSCH, E; WEISS-SCHNEEWEISS, H. Molecular and cytogenetic evidence for an allotetraploid origin of *Chenopodium quinoa* and *C. berlandieri* (Amaranthaceae). *Molecular Phylogenetics and Evolution*. 2016, vol 100, p. 109-123.

- KOZIOL, M. J. Quinoa: a potential new oil crop. En: JANICK, J. y SIMON, J. E. (edit). *New Crops*. New York: Wiley, 1993, p. 328-336. ISBN 0-471-59374-5.
- LECLAIR, B; FRÉGEAU, C. J; BOWEN, K. L. y FOURNEY, R. M. Systematic analysis of stutter percentages and allele peak height and peak area ratios at heterozygous STR loci for forensic casework and database samples. *Journal of Forensic Sciences*. 2004, vol 49, num 5, p. 1-13.
- LESCANO, R. J. L. *Cultivo de la quinua*. Puno, Perú: Universidad Nacional Técnica del Altiplano, Centro de Investigaciones en Cultivos Andinos, 1981.
- LITT, M. y LUTY, J. A. A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *American Journal of Human Genetics*. 1989, vol 44, num 3, p. 397-401.
- MADESIS, P; GANOPOULOS, I. y TSAFTARIS, A. Microsatellites: Evolution and Contribution. En: KANTARTZI, S. K. (edit). *Microsatellites: Methods and Protocols*. Illinois, EE. UU: Humana Press, 2013, p. 1-13. ISBN 978-1-62703-389-3.
- MASON, S. L; STEVENS, M. R; JELLEN, E. N; BONIFACIO, A; FAIRBANKS, D. J; COLEMAN, C. E; MCCARTY, R. R; RASMUSSEN, A. G. y MAUGHAN, P. J. Development and use of microsatellites markers for germplasm characterization in quinoa (*Chenopodium quinoa* Willd.). *Crop Science*. 2005, vol 45, num 5, p. 1618-1630.
- MAUGHAN, P. J; BONIFACIO, A; JELLEN, E. N; STEVENS, M. R; COLEMAN, C. E; RICKS, M; MASON, S. L; JARVIS, D. E; GARDUNIA, B. W. y FAIRBANKS, D. J. A genetic linkage map of quinoa (*Chenopodium quinoa*) based on AFLP, RADP, and SSR markers. *Theoretical and Applied Genetics*. 2004, vol 109, num 6, p. 1188-1195.
- MAUGHAN, P. J; KOLANO, B. A; MALUSZYNSKA, J; COLES, N. D; BONIFACIO, A; ROJAS, J; COLEMAN, C. E; STEVENS, M. R; FAIRBANKS, D. J; PARKINSON, S. E. y JELLEN, E. N. Molecular and cytological

- characterization of ribosomal DNAs in *Chenopodium quinoa* and *Chenopodium berlandieri*. *Genome*. 2006, vol 49, num 7, p. 825-839.
- MAUGHAN, P. J; SMITH, S. M; ROJAS-BELTRÁN, J. A; ELZINGA, D; RANEY, J. A; JELLEN, E. N; BONIFACIO, A; UDALL, J. A. y FAIRBANKS, D. J. Single nucleotide polymorphism identification, characterization, and linkage mapping in Quinoa. *The Plant Genome*. 2012, vol 5, num 3, p. 114-125.
- MEIRMANS, P. G. y VAN TIENDEREN, P. H. The effects of inheritance in tetraploids on genetic diversity and population divergence. *Heredity*. 2013, vol 110, num 2, p. 131-137.
- MOHAMMADI, S. A. y PRASANNA, B. M. Analysis of genetic diversity in crop plants – salient statistical tools and considerations. *Crop Science*. 2003, vol 43, num 4, p. 1235–1248.
- MUJICA, A. Andean grains and legumes. En: HERNANDO-BERMUJO, J. E. y LEON, J (edit). *Neglected Crops: 1492 from a different perspective Vol. 26*. Roma, Italia: FAO, 1994, p. 131-148. ISBN: 9251032173.
- MUJICA, A; MARCA, S. y JACOBSEN, S. E. Current production and potential of quinoa (*Chenopodium quinoa* Willd.) in Peru. *Food Reviews International*. 2003, vol 19, num 1-2, p. 149-154.
- MUJICA, A. La quínoa indígena, características e historias. En: SEPÚLVEDA, J; THOMET, M. I; PALAZUELOS, F. y MUJICA, A. (edit.). *La kinwa mapuche, recuperación de un cultivo para la alimentación*. Chile: Fundación para la innovación agraria, Ministerio de Agricultura Temuco, 2004, p. 22-42.
- NEI, M. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*. 1978, vol 89, num 3, p. 583-590.
- NEI, M. y CHESSER, R. K. Estimation of fixation índices and gene diversities. *Annals of Human Genetics*. 1983, vol 47, num 3, p. 253-259.
- NEI, M. *Molecular Evolutionary Genetics*. 1era ed. New York, EE. UU: Columbia University Press, 1987, 512 p. ISBN: 0231063210.

- OBBARD, D. J; HARRIS, S. A. y PANNELL, J. R. Simple allelic-phenotype diversity and differentiation statistics for allopolyploids. *Heredity*. 2006, vol 97, num 4, p. 296-303.
- OSATINSKY, R. ¿Qué es la electroforesis capilar? *Bioquímica y Patología Clínica*. 2007, vol 71, num 2, p. 60-66.
- OSAWARU, M. E; OGWU, M. C. y AIWANSOBA, R. O. Hierarchical approaches to the analysis of genetic diversity in plants: a systematic overview. *University of Mauritius Research Journal*. 2015, vol 21, p. 1-36.
- OSHODI, A; OGUNGBENLE, H. y OLADIMEJI, M. Chemical composition, nutritionally valuable minerals and functional properties of benniseed, pearl millet and quinoa flours. *International Journal of Food Sciences and Nutrition*. 1999, vol 50, num 5, p. 325-331.
- OTT, J. Strategies for characterizing highly polymorphic markers in human gene mapping. *American Journal of Human Genetics*. 1992, vol 51, num 2, p. 283-290.
- PALMER, M. W. Ordination Methods – an overview [en línea]. Oklahoma: Botany Department, Oklahoma State University, 2004 [ref. del 6 de julio del 2017]. Disponible en web: < <http://ordination.okstate.edu/overview.htm>>.
- PALOMINO, G. H; SEGURA, M. D; BYE, R. B. y MERCADO, R. P. Cytogenetic distinction between *Teloxys* and *Chenopodium* (Chenopodiaceae). *The Southwestern Naturalist*. 1990, vol 35, num 3, p. 351-353.
- PARASNIS, A. S; RAMAKRISHNA, W; CHOWDARI, K. V; GUPTA, V. S. y RANJEKAR, P. K. Microsatellite (GATA)_n reveals sex-specific differences in Papaya. *Theoretical and Applied Genetics*. 1999, vol 99, num 6, p. 1047-1052.
- PLANELLA, M. T; LÓPEZ, M. L. y BRUNO M. C. La domesticación y distribución prehistórica. En: BAZILE, D. et al. (edit). *Estado del arte de la quinua en el mundo en 2013*. Montpellier, Francia: FAO (Santiago de Chile) y CIRAD, 2014, p. 33-48. ISBN: 978-92-5-308558-3.

- PODANI, J. Introduction to the exploration of multivariate biological data. Backhuys Publishers, 2000, 407 p. ISBN: 9057820676.
- PRATT, C. AFLP analysis of genetic diversity in the USDA *Chenopodium quinoa* collection. Tesis. Universidad Brigham Young, Provo, Utah, EE. UU, 2003.
- PRITCHARD, J. K; STEPHENS, M. y DONNELLY, P. Inference of Population Structure Using Multilocus Genotype Data. *Genetics*. 2000, vol 155, num 2, p. 945–959.
- PRITCHARD, J. K; WEN, X. y FALUSH, D. *Documentation for structure software: Version 2.3*. Chicago, EE. UU: Departamento de Genética Humana, Universidad de Chicago, 2010, 38 p.
- RANHOTRA, G; GELROTH, J; GLASER, B; LORENZ, K. y JOHNSON, D. Composition and protein nutritional quality of quinoa. *Cereal Chemistry*. 1993, vol 70, num 3, p. 303-305.
- R DEVELOPMENT CORE TEAM. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2008. ISBN: 3-900051-07-0
- REIF, J. C; MELCHINGER, A. E. y FRISCH, M. Genetical and mathematical properties of similarity and dissimilarity coefficients applied in plant breeding and seed bank management. *Crop Science*. 2005, vol 45, num 1, p. 1-7.
- RETTIG, J. H; WILSON, H. D. y MANHART, J. R. Phylogeny of the Caryophyllales: gene sequence data. *Taxon*. 1992, vol 41, num 2, p. 201-209.
- RISI, J. y GALWEY, N. W. The *Chenopodium* grains of the Andes: Inca crops for modern agriculture. En: COAKER, T. H (edit). *Advances in Applied Biology Vol. 10*. London: Academic Press, 1984, p. 145-216. ISBN-13: 978-0120409105.
- ROJAS, W. Análisis de la diversidad genética del germoplasma de quinua (*Chenopodium quinoa* Willd.) de Bolivia, mediante métodos multivariados. Tesis de Maestría. Universidad Austral de Chile, Facultad de Ciencias Agrarias, Valdivia, Chile, 1998.

- ROJAS, W; PINTO, M; ALANOCA, C; GÓMEZ, L; LEÓN-LOBOS, P; ALERCIA, A; DIULGHEROFF, S; PADULOSI, S. y BAZILE, D. Estado de la conservación ex situ de los recursos genéticos de quinua. En: BAZILE, D. et al. (edit). *Estado del arte de la quinua en el mundo en 2013*. Montpellier, Francia: FAO (Santiago de Chile) y CIRAD, 2014, p. 65-94. ISBN: 978-92-5-308558-3.
- ROULLIER, C; KAMBOUO, R; PAOFA, J; MCKEY, D. y LEBOT, V. On the origin of sweet potato (*Ipomoea batatas* (L.) Lam.) genetic diversity in New Guinea, a secondary centre of diversity. *Heredity*. 2013, vol 110, num 6, p. 594-604.
- RUALES, J. y NAIR, B. M. Nutritional quality of the protein in quinoa (*Chenopodium quinoa* Willd.) seeds. *Plant Foods for Human nutrition*. 1992, vol 42, num 1, p. 1-12.
- SARAVIA, R. La androesterilidad en quinua y forma de herencia. Tesis. Universidad Mayor de San Simón, Cochabamba, Bolivia, 1991.
- SCHUELKE, M. An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology*. 2000, vol 18, num 2, p. 233-234.
- SELKOE, K. A. y TOONEN, R. J. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters*. 2006, vol 9, num 5, p. 615-629.
- SILVESTRI, V. y GIL, F. Alogamia en quinua. Tasa en Mendoza (Argentina). *Revista de la Facultad de Ciencias Agrarias (Universidad Nacional de Cuyo)*. 2000, vol 32, num 1, p. 71-76.
- SIMMONDS, N. W. The breeding system of *Chenopodium quinoa*. I. Male sterility. *Heredity*. 1971, vol 27, p. 73-82.
- SINGH, S. Understanding the weedy *Chenopodium* complex in the north central States. Tesis Doctoral. Universidad de Illinois, Urbana-Champaign, Illinois, 2010.
- STIFT, M; BERENOS, C; KUPERUS, P. y VAN TIENDEREN, P. H. Segregations models for disomic, tetrasomic and intermediate inheritance in tetraploids: A general

- procedure applied to *Rorippa* (Yellow Cress) microsatellite data. *Genetics*. 2008, vol 179, num 4, p. 2113-2123.
- TAPIA, M. Historia y distribución geográfica. En: TAPIA, M (edit). *Quinoa y Kañiwa: cultivos andinos*. Bogotá, Colombia: CIID, Oficina regional para la América Latina, 1979, p. 11-19. ISBN: 0-88936-200-9.
- TAPIA, M; MUJICA, S. A. y CANAHUA, A. Origen, distribución geográfica, y sistemas de producción en quinua. En: Centro de Investigación para el desarrollo. *Primera Reunión sobre Genética y Fijomejoramiento de la Quinoa*. Puno, Perú: Universidad Nacional Técnica del Altiplano, Instituto Boliviano de Tecnología Agropecuaria, Instituto Interamericana de Ciencias Agrícolas, 1980, p. A1-A8.
- TAUTZ, D. Notes on the definition and nomenclature of tandemly repetitive DNA sequences. En: PENA, S. D. J; CHAKRABORTY, R; EPPLEN, J. T. y JEFFREYS, A. J. (edit.). *DNA fingerprinting: state of the science*. Basilea, Suiza: Birkhäuser Basel, 1993, p. 21-28. ISBN: 978-3-0348-8583-6.
- VERAMENDI, S; BONIFACIO, A; CADIMA, X. y ROJAS, W. Caracterización de la diversidad genética de la colección boliviana de quinua, utilizando microsatélites. En: VARGAS, M. (edit). *Congreso científico de la quinua (memorias)*. La Paz, Bolivia, 2013, p. 43-53.
- VIA Y RADA, R. N. Determinación de la diversidad genética de 172 accesiones de la colección nacional de *Chenopodium quinoa* Willd. “quinua” mediante marcadores microsatélites. Tesis Título Profesional. Escuela Profesional de Biología, Facultad de Ciencias Biológicas, Universidad Ricardo Palma, Lima, Perú, 2015.
- WANG, M. L; Barkley, N. A. y Jenkins, T. M. Microsatellite Markers in Plants and Insects Part I: Applications of Biotechnology. *Genes, Genomes and Genomics*. 2009, vol 3, num 1, p. 54-67.

- WARD, S. M. Response to selection for reduced grain saponin content in quinoa (*Chenopodium quinoa* Willd.). *Field Crops Research*. 2000a, vol 68, num 2, p. 157-163.
- WARD, S. M. Allotetraploid segregation for single-gene morphological characters in quinoa (*Chenopodium quinoa* Willd.). *Euphytica*. 2000b, vol 116, num 1, p. 11-16.
- WEBER, J. L. Human DNA polymorphisms and methods of analysis. *Current Opinion in Biotechnology*. 1990, vol 1, num 2, p. 166-171.
- WILFINGER, W. W; MACKEY, K. y CHOMCZYNSKI, P. Effect of pH and ionic strength on the spectrophotometric assessment of nucleic acid purity. *Biotechniques*. 1997, vol 22, num 3, p. 478-481.
- WILSON, H. D. Genetic control and distribution of leucine aminopeptidase in the cultivated chenopods (Chenopodiaceae) and related weed taxa. *Biochemical Genetics*. 1976, vol 14, num 11, p. 913-919.
- WILSON, H. D. Artificial hybridization among species of *Chenopodium* sect. *Chenopodium*. *Systematic Botany*. 1980, vol 5, num 3, p. 253-263.
- WILSON, H. D. Quinoa systematics I: domesticated populations. *Economic Botany*. 1988, vol 42, num 4, p. 461-477.
- WILSON, H. D. Quinoa and relatives (*Chenopodium* sect. *Chenopodium* subsect. *Cellulata*). *Economic Botany*. 1990, vol 44, num 3, p. 92-110.
- WRIGHT, S. The genetical structure of populations. *Annals of Human Genetics*. 1950, vol 15, num 1, p. 323-354.
- YAÑEZ, V. O. Aislamiento y caracterización de marcadores moleculares microsatélites a partir de la construcción de librerías genómicas enriquecidas de camote (*Ipomoea batatas* (L.) Lam.). Tesis Título Profesional. Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Biológicas, EAP de Ciencias Biológicas, Lima, Perú, 2002.

- ZANE, L; BARGELLONI, L. y PATARNELLO, T. Strategies for microsatellite isolation: a review. *Molecular Ecology*. 2002, vol 11, num 1, p. 1-16.
- ZHONG, S. y GHOSH, J. A unified framework for model-based clustering. *Journal of Machine Learning Research*. 2003, vol 4, p. 1001-1037.

11. GLOSARIO

Accesión: Una muestra distinta, singularmente identificable de semillas que representa un cultivar, una línea de cría o una población y que se mantiene almacenada para su conservación y uso (FAO, 2014).

Alelo nulo: Un alelo que falla en amplificar utilizando cebadores específicos de un locus y que no es observado debido a la baja calidad de la muestra (Dufresne et al., 2014).

Alelo privado: Alelo que se encuentra solamente en una población o región, también llamado alelo exclusivo o único.

Alopoliploide: Organismo poliploide que se origina a partir de una duplicación del genoma después de una hibridación, de tal manera que existe dos conjuntos de homeólogos del mismo cromosoma. El dogma consiste en que generalmente presenta herencia disómica debido a un apareamiento preferencial entre cromosomas de un mismo genoma ancestral. Sin embargo, es posible una herencia polisómica al menos en algunos loci o regiones cromosómicas (Dufresne et al., 2014).

Autogamia: Proceso de fertilización de gametos de una misma flor en una misma planta.

Autopoliploide: Organismo poliploide originado a partir de una duplicación del genoma dentro de una especie de tal manera que todas las variantes de un mismo cromosoma son homólogos. El dogma es que generalmente presenta herencia polisómica debido a que no hay un apareamiento preferencial entre cromosomas homólogos. Sin embargo, como la duplicación de un genoma conlleva inevitablemente a una divergencia, se predice una herencia disómica con el tiempo (Dufresne et al., 2014).

Banco de germoplasma: Centro para la conservación de recursos genéticos en condiciones adecuadas para prolongar sus vidas (FAO, 2014).

Caracterización: Registro de caracteres altamente heredables que pueden ser vistos fácilmente y que se expresan en todos los ambientes (FAO, 2014).

Cebador: Una corta secuencia de nucleótidos que se aparea con una hebra de ADN y que provee un extremo libre en donde la enzima ADN polimerasa comienza la síntesis del segmento complementario de la hebra de ADN (Frankham et al., 2002).

Colección: Grupo de accesiones de germoplasma mantenido para un propósito específico en condiciones definidas (FAO, 2014).

Conservación *ex situ*: Conservación de la diversidad biológica fuera de su hábitat natural, por ejemplo, en bancos de germoplasma o bancos de semillas (FAO, 2014).

Datos de pasaporte: Información básica sobre el origen de una accesión, como los detalles registrados en el lugar de recolección, la información pertinente sobre el pedigrí o de otro tipo que ayude a identificar una accesión (FAO, 2014).

Dosis alélica: Número de copias de cada alelo en un locus determinado en un genotipo poliploide (Dufresne et al., 2014).

Electroferograma: Gráfico del tiempo (eje X) proporcional al tamaño de la molécula marcada con un fluoróforo versus el nivel de fluorescencia emitida (eje Y) por tal molécula marcada. Es el resultado gráfico de una electroforesis capilar.

Equilibrio Hardy-Weinberg: El equilibrio de las frecuencias genotípicas alcanzado en una población de apareamiento aleatorio sin la intervención de una fuerza que perturbe el equilibrio como la mutación, migración, selección o deriva (Frankham et al., 2002).

Germoplasma: Material genético que constituye la base física de la herencia y que se transmite de una generación a la sucesiva mediante las células germinales (FAO, 2014).

Herencia disómica: Tipo de herencia típico de aloploidos debido a un apareamiento preferencial entre cromosomas derivados de la misma especie ancestral. Esto

significa que los alelos derivados de la misma especie ancestral se segregan como en un organismo diploide, de manera que la descendencia recibe una copia de cada padre. Por lo tanto, no se espera la recombinación entre copias derivadas de ancestros diferentes (homeólogos) (Dufresne et al., 2014).

Herencia polisómica: Tipo de herencia típico de autoploidoides, donde todas las variantes del mismo cromosoma pueden aparearse en meiosis. Esto significa que los alelos se combinarán en los gametos en todas las posibles combinaciones (Dufresne et al., 2014).

Heterocigoto parcial: Genotipo poliploide en donde no se sabe con certeza el número de copias de cada alelo aún si la herencia del locus es codominante (como microsatélites), por ejemplo, en tetraploides los genotipos pueden ser homocigotos (AAAA, BBBB, CCCC), heterocigotos completos (ABCD, ABCF) o heterocigotos parciales (ABBC, AABC o ABCC, en un electroferograma o gel no se pueden diferenciar entre los tres posibles genotipos).

Homeólogo: Loci o cromosomas divergentes en genomas aloploidoides que usualmente no se aparean durante la meiosis debido a que se derivan de distintas líneas parentales (Dufresne et al., 2014).

Homólogo: Loci o cromosomas que usualmente se aparean durante la meiosis debido a que derivan de una misma línea parental (Dufresne et al., 2014).

Motivo: Secuencia de nucleótidos con determinadas bases nitrogenadas que se repiten en tándem en una sección altamente repetitiva del ADN como es el caso de las regiones microsatélites, ejemplo: motivos dinucleótidos como CA, AT, CG o trinucleótidos como AAT, ACA, CGC.

Muestra: Parte de una población que se utiliza para estimar las características del conjunto (FAO, 2014).

Pico: Representación gráfica de una hebra de ADN marcada fluorescentemente en un electroferograma cuya altura es proporcional a la cantidad de ADN marcado.

Pico tartamudo: Artefactos debido al deslizamiento de la polimerasa durante la PCR de secuencias altamente repetitivas (microsatélites), visibles como uno o mas picos de longitud mas corta o larga que la longitud del alelo (Dufresne et al., 2014).

Reacción en cadena de la polimerasa: Metodo usado para hacer multiples copias de un segmento especifico de ADN utilizando un par de segmentos que flanquean el segmento de interés (Frankham et al., 2002).

Tasa de tetrasomía (τ): Parámetro que indica la proporción de gametos formados por asociaciones cromosómicas meióticas aleatorias (apareamiento de cuatrivalentes o bivalentes al azar) y puede tomar valores de 0 (totalmente disómico) a 1 (totalmente tetrasómico) (Stift et al., 2008).

Variedad: División reconocida de una especie, que sigue en rango por debajo de la subespecie. Se distingue por características como el color de las flores, el color de las hojas, y el tamaño de la planta madura. El término se considera sinónimo de cultivar (FAO, 2014).

12. ANEXOS

12.1. Tablas

Tabla 5.1. Datos pasaporte de las 129 accesiones de *Chenopodium quinoa* Willd. del banco de germoplasma de la EEA Illpa del INIA.

N°	Código Nacional INIA	Departamento	Provincia	Distrito	Localidad	Latitud	Longitud
1	PER004275	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
2	PER004276	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
3	PER004277	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
4	PER004278	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
5	PER004279	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
6	PER004280	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
7	PER004281	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
8	PER004282	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
9	PER004283	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
10	PER004284	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
11	PER004285	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
12	PER004286	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
13	PER004287	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
14	PER004288	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
15	PER004289	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
16	PER004290	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
17	PER004291	Puno	Puno	Puno	UNTA	-15.8392	-70.02917

18	PER004292	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
19	PER004293	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
20	PER004295	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
21	PER004296	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
22	PER004297	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
23	PER004298	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
24	PER004299	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
25	PER004300	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
26	PER004301	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
27	PER004302	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
28	PER004303	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
29	PER004304	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
30	PER004305	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
31	PER004306	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
32	PER004307	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
33	PER004308	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
34	PER004309	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
35	PER004310	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
36	PER004311	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
37	PER004312	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
38	PER004313	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
39	PER004314	Puno	Puno	Puno	UNTA	-15.8392	-70.02917

40	PER004315	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
41	PER004316	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
42	PER004317	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
43	PER004318	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
44	PER004319	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
45	PER004320	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
46	PER004321	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
47	PER004322	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
48	PER004323	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
49	PER004324	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
50	PER004325	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
51	PER004326	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
52	PER004327	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
53	PER004328	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
54	PER004329	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
55	PER004330	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
56	PER004331	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
57	PER004333	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
58	PER004334	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
59	PER004335	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
60	PER004336	Puno	Puno	Puno	UNTA	-15.8392	-70.02917
61	PER004499	Puno	El Collao	Ilave	Ullacachi	-16.16853	-69.7088

62	PER004500	Puno	San Román	Cabana	Puca chupa	-15.62223	-70.32474
63	PER004501	Puno	Puno	Acora	Anccacca	-16.1404	-69.71057
64	PER004502	Puno	Puno	Acora	Cusini	-16.17361	-69.74469
65	PER004503	Puno	Puno	Acora	Quenafaja	-16.1262	-69.7029
66	PER004504	Puno	Puno	Acora	Laccachi	-16.18611	-69.78653
67	PER004505	Puno	Puno	Acora	Quenafaja	-16.1262	-69.7029
68	PER004506	Puno	Puno	Acora	Canchon	-15.97415	-69.79878
69	PER004507	Puno	Puno	Pichacani	Oqe huichinca	-16.10439	-70.13294
70	PER004508	Puno	Puno	Tiquillaca	Tiquillaca	-15.79779	-70.18548
71	PER004509	Puno	Puno	Pichacani	Oqe huichinca	-16.10439	-70.13294
72	PER004510	Puno	Puno	Vilque	Kere - vilque	-15.76652	-70.25628
73	PER004511	Puno	Puno	Pichacani	Oqe huichinca	-16.10439	-70.13294
74	PER004512	Puno	Puno	Vilque	Sector queta	-15.76652	-70.25628
75	PER004513	Puno	Puno	Acora	Amparani	-16.11711	-69.70585
76	PER004514	Puno	San Román	Cabana	Puca chupa	-15.62223	-70.32474
77	PER004515	Puno	Puno	Acora	Quenafaja	-16.1262	-69.7029
78	PER004516	Puno	Puno	Acora	Cusini	-16.17361	-69.74469
79	PER004517	Puno	Puno	Tiquillaca	Kere (parc.)	-15.79779	-70.18548
80	PER004518	Puno	Puno	Vilque	Kere - vilque	-15.76652	-70.25628
81	PER004519	Puno	Puno	Acora	Quelca opojani	-15.97415	-69.79878
82	PER004527	Puno	Puno	Acora	Laccachi	-16.18611	-69.78653
83	PER004528	Puno	Puno	Paucarcolla	Anexo Illpa	-15.71244	-70.08307

84	PER004529	Puno	Puno	Paucarcolla	Anexo Illpa	-15.71244	-70.08307
85	PER004530	Puno	Puno	Paucarcolla	Anexo Illpa	-15.71244	-70.08307
86	PER004531	Puno	Puno	Paucarcolla	Anexo Illpa	-15.71244	-70.08307
87	PER004532	Puno	Yunguyo	Yunguyo	Tahuaco	-16.31286	-69.06543
88	PER004533	Puno	Yunguyo	Yunguyo	Tahuaco	-16.31286	-69.06543
89	PER004534	Puno	Yunguyo	Yunguyo	Tahuaco	-16.31286	-69.06543
90	PER004535	Puno	Yunguyo	Yunguyo	Tahuaco	-16.31286	-69.06543
91	PER004536	Puno	Yunguyo	Yunguyo	Tahuaco	-16.31286	-69.06543
92	PER004537	Puno	Yunguyo	Yunguyo	Tahuaco	-16.31286	-69.06543
93	PER004538	Puno	Yunguyo	Yunguyo	Tahuaco	-16.31286	-69.06543
94	PER004539	Puno	Yunguyo	Yunguyo	Tahuaco	-16.31286	-69.06543
95	PER004540	Puno	Puno	Paucarcolla	Anexo Illpa	-15.71244	-70.08307
96	PER004550	Puno	Chucuito	Pomata	Japurani	-16.27294	-69.29308
97	PER004551	Puno	Chucuito	Pomata	Condorccollo	-16.27294	-69.29308
98	PER004552	Puno	Chucuito	Juli	Achucollani	-16.32052	-69.55613
99	PER004553	Puno	Chucuito	Juli	Achucollani	-16.32052	-69.55613
100	PER004554	Puno	Chucuito	Pomata	P. Central	-16.27294	-69.29308
101	PER004555	Puno	Chucuito	Pomata	Condorccollo	-16.27294	-69.29308
102	PER004556	Puno	Chucuito	Pomata	Tucumpalla	-16.27294	-69.29308
103	PER004557	Puno	Chucuito	Pomata	Huariphujo	-16.27294	-69.29308
104	PER004558	Puno	Chucuito	Huacullani	Totorama a	-16.6947	-69.375
105	PER004559	Puno	Chucuito	Juli	Achucollani	-16.32052	-69.55613

106	PER004560	Puno	Chucuito	Huacullani	Totorama b	-16.6947	-69.375
107	PER004561	Puno	Chucuito	Huacullani	Totorama a	-16.6947	-69.375
108	PER004562	Puno	Chucuito	Pomata	Huariphujo	-16.27294	-69.29308
109	PER004563	Puno	Chucuito	Pomata	Huariphujo	-16.27294	-69.29308
110	PER004564	Puno	Chucuito	Pisacoma	Chiaramaya	-16.9042	-69.3658
111	PER004565	Puno	Chucuito	Pomata	Mejani	-16.31846	-69.27758
112	PER004566	Puno	Chucuito	Huacullani	Totorama a	-16.6947	-69.375
113	PER004567	Puno	Chucuito	Pisacoma	Chiaramaya	-16.9042	-69.3658
114	PER004568	Puno	Chucuito	Pomata	Tucumpalla	-16.27294	-69.29308
115	PER004569	Puno	Chucuito	Huacullani	Totorama b	-16.6947	-69.375
116	PER004570	Puno	Chucuito	Pisacoma	Chiaramaya	-16.9042	-69.3658
117	PER004571	Puno	Chucuito	Pomata	P. Central	-16.27294	-69.29308
118	PER004572	Puno	Chucuito	Pomata	Mejani	-16.31846	-69.27758
119	PER004573	Puno	Chucuito	Pomata	Condorccollo	-16.27294	-69.29308
120	PER004574	Puno	Chucuito	Pomata	Japurani	-16.27294	-69.29308
121	PER004575	Puno	Chucuito	Pomata	Japurani	-16.27294	-69.29308
122	PER004576	Puno	Chucuito	Pomata	Japurani	-16.27294	-69.29308
123	PER004578	Puno	Chucuito	Pomata	Pomata	-16.27294	-69.29308
124	PER004579	Puno	Puno	Paucarcolla	Anexo Illpa	-15.71244	-70.08307
125	PER004581	Puno	Puno	Acora	Marca esqueña	-16.03098	-69.75494
126	PER004582	Puno	Puno	Acora	Marca esqueña	-16.03098	-69.75494
127	PER004583	Puno	Puno	Acora	Marca esqueña	-16.03098	-69.75494

128	PER004584	Puno	Chucuito	Pomata	Huacani	-16.27901	-69.32172
129	PER004585	Puno	Puno	Puno	Ichu	-15.87837	-69.94208

Tabla 5.2. Protocolo de preparación de los tampones de extracción y de carga utilizados en la extracción de ADN genómico de *Chenopodium quinoa* Willd.

Tampón de extracción CTAB 2X		
Solución madre (Stock)	Final	100 ml
Tris-HCl pH 8.0 a 1 M	100 mM	10 ml
CTAB	2%	2 g
NaCl	1.4 M	8.12 g
EDTA pH 8.0 a 0.7 M	28 mM	4 ml
Polivinilpirrolidona (PVP)	2.5%	2.5 g
H ₂ O miliQ		Completar a 100 ml
Tampón de extracción CTAB 10X		
Solución madre (Stock)	Final	100 ml
Tris-HCl pH 8.0 a 1 M	100 mM	10 ml
CTAB	10%	10 g
NaCl	0.7 M	2 g
EDTA pH 8.0 a 0.5 M	10 mM	2 ml
H ₂ O miliQ		Completar a 100 ml
Tampón de carga SALB 10X		
Solución madre (Stock)	Final	100 ml
Azul de bromofenol	0.15%	150 mg
Xilencianol	0.15%	150 mg
Naranja G	0.2%	200 mg
Sucrosa	60%	60 g
TBE 10X	5% (v/v)	5 ml
H ₂ O miliQ		Completar a 100 ml

Tabla 5.3. Características de los 15 marcadores microsatélites elegidos para el estudio de la diversidad genética de *Chenopodium quinoa* Willd. Las porciones en negrita de las secuencias de los cebadores QAAT24, QAAT50, QAAT70, QAAT74, QAAT76 y KGA3 representan nucleótidos (gcttct) añadidos al cebador por Christensen et al. (2007) para facilitar la PCR.

Nombre	Secuencia directa (5'-3')	Secuencia reversa (5'-3')	Motivo de repetición	Temp. de hibridación (°C)	Referencia
QGA021	cacgaaaccaactcctctca	caccacaatcaccaccttg	(CT) ₂₁	59	Mason et al., 2005
QGA024	aatcccgaaagcatattaacca	aggaatggctctctggcaca	(AG) ₂₁	59	Mason et al., 2005
QAAT027	aatgaggaggcaatgcaaag	cggctccctaccaatttctt	(TATT) ₅ (ATT) ₁₃	59	Mason et al., 2005
QAAT112	cccgatccaccataagagaa	tgaagtgtgaagattggagaatgaca	(ATT) ₁₃	59	Mason et al., 2005
QAAT106	tcagtaagataatcccatcagtaag	aaaatcccctctataattaccaa	(TAA) ₉	58	Mason et al., 2005
QGA032	ggtgaattcgaatggcaag	tggcagttggatccactataaa	(AG) ₂₉	59	Mason et al., 2005
KAAT006	tctgcaggatcggtaacctt	ttgtatctcggctcccact	(ATT) ₁₀ (GTT) ₇ (ATT) ₃	54	Jarvis et al., 2008
QAAT24	gcttct accataaacagcacccacctt	agggatcaatctgttcattca	(ATT) ₁₀	54	Mason et al., 2005
QAAT50	ggcacgtgctgctactcata	gcttct atggcgaatggtaatttgc	(ATT) ₁₇	53	Mason et al., 2005
KAAT037	tcaacctccgaatcctatcaa	ggatgctgattggtggataaa	(TAA) ₁₉	53	Jarvis et al., 2008
QAAT70	tgaacaggatcgatcatagtcaa	gcttct cggtcatcatctgaccaat	(ATT) ₁₅	57	Mason et al., 2005
KGA3	attgccgacaatgaacgaat	gcttct atgtaaatggcatgtccaac	(GA) ₁₆	58	Jarvis et al., 2008
QAAT76	gcttcatgtgttataaaatgccaat	gcttct tctcggctcccactaatttt	(ATT) ₃₀	62	Mason et al., 2005
QAAT74	gcttct atggaacacccatccgataa	atgcctatcctcatcctcca	(ATT) ₁₄	60	Mason et al., 2005
QAAT100	ggcatccagaggtcagtcctt	gcaattcttcctaataacaacaacaa	(ACTACC) ₈ ACTGTT (ATTGTT) ₂₃ (ATT) ₅ G TT(ATT) ₂ (ACT) ₁₁	60	Mason et al., 2005

Tabla 5.4. Protocolo de preparación de reactivos para PCR con un solo par de cebadores (QAAT106, QGA032, KAAT006, QAAT24, QAAT50, KAAT037, QAAT70, KGA03, QAAT76).

Reactivo	Conc. inicial	Conc. final	Volumen (1 reacc.)
NFW	-	-	2.63 ul
Tampón KAPA 5X	5X	1X	2 ul
MgCl ₂	25 mM	2.25 mM	0.9 ul
dNTPs	5 mM	0.2 mM	0.4 ul
Directo	1 uM	0.04 uM	0.4 ul
Reverso	1 uM	0.06 uM	0.6 ul
M13	1 uM	0.06 uM	0.6 ul
BSA	10 mg/ml	0.4 mg/ml	0.4 ul
Taq Kapa	5 U/ul	0.35 U	0.07 ul
DNA	20 ng/ul	-	2 ul
Volumen final	-	-	10 ul

Tabla 5.5. Protocolo de preparación de reactivos para PCR multiplex (2 pares de cebadores por reacción).

Reactivo	QGA021 ⁽¹⁾ y QGA024 ⁽²⁾		QAAT112 ⁽¹⁾ y QAAT027 ⁽²⁾		QAAT100 ⁽¹⁾ y QAAT74 ⁽²⁾	
	Conc. inicial	Conc. final	Conc. inicial	Conc. final	Conc. inicial	Conc. final
NFW	-	-	-	-	-	-
Tampón KAPA 5X	5X	1X	5X	1X	5X	1X
MgCl ₂	25 mM	2.25 mM	25 mM	2.25 mM	25 mM	2.25 mM
dNTPs	5 mM	0.2 mM	5 mM	0.2 mM	5 mM	0.2 mM
Directo ⁽¹⁾	1 uM	0.06 uM	1 uM	0.08 uM	1 uM	0.02 uM
Reverso ⁽¹⁾	1 uM	0.08 uM	1 uM	0.1 uM	1 uM	0.08 uM
Directo ⁽²⁾	1 uM	0.04 uM	1 uM	0.06 uM	1 uM	0.025 uM
Reverso ⁽²⁾	1 uM	0.06 uM	1 uM	0.08 uM	1 uM	0.1 uM
M13	1 uM	0.06 uM	1 uM	0.08 uM	1 uM	0.1 uM
BSA	10 mg/ml	0.4 mg/ml	10 mg/ml	0.4 mg/ml	10 mg/ml	0.4 mg/ml
Taq Kapa	5 U/ul	0.35 U	5 U/ul	0.35 U	5 U/ul	0.35 U
DNA	20 ng/ul	-	20 ng/ul	-	20 ng/ul	-
Volumen final	-	10 ul	-	10 ul	-	10 ul

Tabla 5.6. Condiciones de la corrida electroforética utilizando el analizador genético ABI 3130xl.

Parámetro	Valor
Temperatura de horno	60 °C
Voltaje de precorrida	15 °C
Tiempo de precorrida	180 segundos
Voltaje de inyección	1.2 kvoltios
Tiempo de inyección	50 segundos
Voltaje de corrida	15 kvoltios
Tiempo de corrida por inyección	1140 segundos

Tabla 5.7. Organización de los marcadores microsatélites en grupos para la electroforesis capilar y sus rangos de tamaño alélico respectivos (en pb). El rango de tamaño no incluye los 19 nucleótidos del cebador M13F (-29). Los colores azul, verde, amarillo y rojo corresponden al tipo de fluorescencia emitida por los fluoróforos 6-FAM, VIC[®], NED[®] y PET[®] respectivamente.

Marcador		Rango de tamaño (pb)	Fluoróforo
PRIMER GRUPO	QGA021	141-161	6-FAM
	QGA024	166-221	
	QAAT027	140-176	VIC [®]
	QAAT112	176-221	
	QAAT106	251-341	NED [®]
	QGA032	151-206	PET [®]
SEGUNDO GRUPO	KAAT006	171-206	6-FAM
	QAAT24	186-241	VIC [®]
	QAAT50	181-251	NED [®]
	KAAT037	191-331	PET [®]
TERCER GRUPO	QAAT70	146-191	6-FAM
	KGA3	141-171	VIC [®]
	QAAT76	131-216	NED [®]
	QAAT74	151-241	PET [®]
	QAAT100	281-441	

Tabla 6.1. Índice de pureza 260/280 y concentración de las soluciones de ADN genómico extraído.

N°	Accesión	Índice 260/280	Conc. (ng/μL)	N°	Accesión	Índice 260/280	Conc. (ng/μL)
1	PER004275	2.026	185.479	66	PER004504	2.131	102.329
2	PER004276	2.021	328.745	67	PER004505	2.133	233.473
3	PER004277	2.262	279.797	68	PER004506	2.168	154.895
4	PER004278	2.061	284.384	69	PER004507	2.124	264.913
5	PER004279	2.19	325.051	70	PER004508	2.193	197.595
6	PER004280	2.117	217.228	71	PER004509	2.049	69.139
7	PER004281	2.216	353.519	72	PER004510	2.045	160.134
8	PER004282	2.216	195.707	73	PER004511	2.019	238.845
9	PER004283	2.043	204.746	74	PER004512	2.069	240.849
10	PER004284	2.128	226.918	75	PER004513	2.062	107.259
11	PER004285	2.037	211.896	76	PER004514	2.051	142.028
12	PER004286	2.014	130.678	77	PER004515	2.062	99.239
13	PER004287	2.108	227.068	78	PER004516	2.148	151.197
14	PER004288	2.162	176.118	79	PER004517	2.084	137.569
15	PER004289	2.096	237.827	80	PER004518	2.046	143.158
16	PER004290	1.935	215.196	81	PER004519	2.114	242.144
17	PER004291	2.064	167.685	82	PER004527	2.164	325.814
18	PER004292	2.061	146.514	83	PER004528	2.017	196.578
19	PER004293	2.117	193.036	84	PER004529	2.056	272.684
20	PER004295	2.083	156.77	85	PER004530	2.057	415.661
21	PER004296	2.197	151.98	86	PER004531	2.099	215.796
22	PER004297	2.085	187.679	87	PER004532	2.103	387.989
23	PER004298	2.168	126.955	88	PER004533	2.086	318.271
24	PER004299	2.133	188.757	89	PER004534	2.087	459.835
25	PER004300	1.885	82.024	90	PER004535	2.106	374.923
26	PER004301	2.037	182.224	91	PER004536	2.142	246.345
27	PER004302	2.089	158.836	92	PER004537	2.187	223.547
28	PER004303	2.044	171.183	93	PER004538	1.92	144.881
29	PER004304	1.985	131.854	94	PER004539	2.13	144.183
30	PER004305	2.254	165.817	95	PER004540	2.088	151.857
31	PER004306	2.17	334.148	96	PER004550	2.125	204.263
32	PER004307	2.144	219.412	97	PER004551	2.166	178.981
33	PER004308	2.039	235.492	98	PER004552	2.184	246.114
34	PER004309	2.067	178.036	99	PER004553	2.166	224.478
35	PER004310	1.988	97.157	100	PER004554	2.197	201.137
36	PER004311	2.108	244.781	101	PER004555	2.112	236.298
37	PER004312	2.167	210.035	102	PER004556	2.165	262.852
38	PER004313	2.191	347.558	103	PER004557	2.187	241.55
39	PER004314	2.19	211.283	104	PER004558	2.173	241.987

40	PER004315	2.185	212.923	105	PER004559	2.084	182.589
41	PER004316	2.313	296.459	106	PER004560	2.04	137.792
42	PER004317	2.178	403.592	107	PER004561	2.139	177.28
43	PER004318	2.216	310.454	108	PER004562	2.085	236.51
44	PER004319	2.178	201.402	109	PER004563	2.099	208.784
45	PER004320	2.168	338.987	110	PER004564	1.941	230.295
46	PER004321	2.205	326.702	111	PER004565	2.017	67.978
47	PER004322	2.185	287.161	112	PER004566	1.989	276.99
48	PER004323	2.13	121.687	113	PER004567	2.187	171.389
49	PER004324	2.125	238.122	114	PER004568	1.949	114.28
50	PER004325	2.2	373.977	115	PER004569	2.133	107.61
51	PER004326	2.072	261.172	116	PER004570	2.042	152.048
52	PER004327	2.195	250.872	117	PER004571	1.955	586.829
53	PER004328	2.076	204.193	118	PER004572	1.956	271.381
54	PER004329	2.167	281.321	119	PER004573	2.135	203.565
55	PER004330	2.106	193.003	120	PER004574	2.15	104.112
56	PER004331	2.089	112.529	121	PER004575	1.982	153.329
57	PER004333	2.17	116.561	122	PER004576	2.113	341.611
58	PER004334	2.141	170.119	123	PER004578	1.938	153.524
59	PER004335	2.089	178.739	124	PER004579	2.114	171.037
60	PER004336	2.149	397.748	125	PER004581	1.988	241.021
61	PER004499	2.148	299.175	126	PER004582	2.109	250.102
62	PER004500	2.151	256.78	127	PER004583	2.04	112.01
63	PER004501	2.148	309.391	128	PER004584	2.016	152.772
64	PER004502	2.042	223.547	129	PER004585	1.951	276.822
65	PER004503	2.092	111.974				

Tabla 6.2. Promedio de la probabilidad logarítmica de los datos ($\text{LnP}(K)$) y del valor ΔK (Evanno et al., 2005) para los valores K del 1 al 7 de la segunda evaluación (20 repeticiones, $K = 1-7$). La tabla se generó utilizando el programa Structure Harvester Web versión 0.6.94. El color amarillo resalta el K con el valor de ΔK más alto.

K	Reps	Mean $\text{LnP}(K)$	Stdev $\text{LnP}(K)$	$\text{Ln}'(K)$	$\text{Ln}''(K)$	Delta K
1	20	-11339.735000	0.668285	—	—	—
2	20	-10921.905000	46.908258	417.830000	4.140000	0.088257
3	20	-10499.935000	19.871463	421.970000	133.890000	6.737803
4	20	-10211.855000	15.243514	288.080000	75.305000	4.940134
5	20	-9999.080000	24.293287	212.775000	13.885000	0.571557
6	20	-9800.190000	15.833838	198.890000	79.545000	5.023735
7	20	-9680.845000	42.535329	119.345000	—	—

Tabla 6.3. Membresía de cada accesión a los grupos hallados por el DAPC según sus probabilidades posteriores. Las accesiones con un asterisco fueron consideradas accesiones mixtas (probabilidad de membresía asociada < 0.8). Los colores rojo, verde y azul corresponden a los grupos K1, K2 y K3 respectivamente.

N°	Código Nacional INIA	Distrito	Grupo asignado por el DAPC	Probabilidades Posteriores		
				1	2	3
1	PER004275	Puno	1	0.999837	0.000001	0.000162
2	PER004276	Puno	1	0.986035	0.004540	0.009424
3	PER004277	Puno	1	0.999181	0.000009	0.000810
4	PER004278*	Puno	2	0.288830	0.710419	0.000751
5	PER004279	Puno	1	0.999997	0.000002	0.000000
6	PER004280*	Puno	2	0.282825	0.717093	0.000082
7	PER004281	Puno	1	0.983323	0.003829	0.012848
8	PER004282	Puno	1	0.999976	0.000013	0.000011
9	PER004283	Puno	1	0.999987	0.000001	0.000012
10	PER004284	Puno	1	0.999143	0.000008	0.000849
11	PER004285	Puno	1	0.980417	0.017455	0.002128
12	PER004286*	Puno	1	0.591782	0.025088	0.383130
13	PER004287	Puno	1	0.999795	0.000186	0.000019
14	PER004288	Puno	1	0.999666	0.000005	0.000329
15	PER004289	Puno	1	0.994783	0.000332	0.004886
16	PER004290	Puno	1	0.999650	0.000006	0.000344
17	PER004291	Puno	3	0.003492	0.000010	0.996497
18	PER004292	Puno	1	0.999865	0.000115	0.000020
19	PER004293	Puno	1	0.802377	0.193191	0.004432
20	PER004295	Puno	2	0.001517	0.924712	0.073771
21	PER004296	Puno	2	0.000013	0.999730	0.000257
22	PER004297	Puno	1	0.974719	0.017361	0.007920
23	PER004298	Puno	1	0.930338	0.067990	0.001672
24	PER004299	Puno	2	0.000379	0.999487	0.000134
25	PER004300	Puno	3	0.000433	0.021329	0.978238
26	PER004301*	Puno	2	0.049107	0.503709	0.447184
27	PER004302	Puno	2	0.000003	0.999990	0.000007
28	PER004303	Puno	2	0.000003	0.999993	0.000004
29	PER004304	Puno	2	0.008182	0.991053	0.000765
30	PER004305	Puno	2	0.006458	0.984306	0.009236
31	PER004306	Puno	2	0.001792	0.996927	0.001281
32	PER004307	Puno	2	0.002311	0.958509	0.039180
33	PER004308	Puno	2	0.000059	0.999935	0.000006
34	PER004309	Puno	3	0.010581	0.032193	0.957226
35	PER004310	Puno	1	0.911172	0.006198	0.082630
36	PER004311	Puno	2	0.012696	0.954059	0.033246

37	PER004312	Puno	2	0.002888	0.974706	0.022405
38	PER004313*	Puno	3	0.231019	0.001434	0.767547
39	PER004314*	Puno	3	0.419441	0.021395	0.559163
40	PER004315	Puno	3	0.000033	0.000002	0.999964
41	PER004316	Puno	3	0.008044	0.002305	0.989652
42	PER004317	Puno	1	0.993885	0.006087	0.000028
43	PER004318	Puno	2	0.000020	0.999742	0.000239
44	PER004319	Puno	1	0.996351	0.001201	0.002447
45	PER004320	Puno	3	0.034876	0.019757	0.945367
46	PER004321	Puno	1	0.891155	0.005623	0.103222
47	PER004322	Puno	1	0.873480	0.020733	0.105787
48	PER004323	Puno	1	0.982869	0.000604	0.016527
49	PER004324*	Puno	3	0.019837	0.409661	0.570502
50	PER004325	Puno	3	0.000198	0.013168	0.986634
51	PER004326	Puno	2	0.003255	0.952325	0.044421
52	PER004327*	Puno	3	0.002528	0.371821	0.625651
53	PER004328	Puno	1	0.981618	0.015844	0.002539
54	PER004329	Puno	3	0.189466	0.002258	0.808277
55	PER004330*	Puno	1	0.628310	0.000266	0.371424
56	PER004331	Puno	3	0.049192	0.043741	0.907067
57	PER004333	Puno	1	0.999958	0.000020	0.000022
58	PER004334*	Puno	2	0.268734	0.698242	0.033024
59	PER004335*	Puno	3	0.483597	0.021709	0.494695
60	PER004336	Puno	3	0.000035	0.002108	0.997857
61	PER004499	Ilave	3	0.000244	0.000469	0.999287
62	PER004500	Cabana	1	0.903928	0.000220	0.095852
63	PER004501	Acora	2	0.015793	0.915623	0.068584
64	PER004502	Acora	2	0.000085	0.996724	0.003191
65	PER004503	Acora	2	0.000003	0.999660	0.000337
66	PER004504*	Acora	3	0.325938	0.000901	0.673160
67	PER004505	Acora	1	0.992872	0.000022	0.007106
68	PER004506	Acora	3	0.000659	0.000008	0.999333
69	PER004507	Pichacani	3	0.003976	0.001316	0.994708
70	PER004508	Tiquillaca	3	0.020729	0.002720	0.976551
71	PER004509	Pichacani	3	0.013948	0.085610	0.900442
72	PER004510	Vilque	3	0.000153	0.000732	0.999116
73	PER004511	Pichacani	3	0.001649	0.000002	0.998349
74	PER004512	Vilque	3	0.000030	0.073776	0.926194
75	PER004513	Acora	2	0.000757	0.998024	0.001219
76	PER004514	Cabana	3	0.019904	0.000478	0.979619
77	PER004515*	Acora	2	0.000089	0.757299	0.242612
78	PER004516	Acora	3	0.029147	0.021200	0.949653
79	PER004517	Tiquillaca	3	0.027399	0.000091	0.972510
80	PER004518	Vilque	3	0.006434	0.002005	0.991561
81	PER004519	Acora	2	0.000008	0.999981	0.000012

82	PER004527	Acora	2	0.000001	0.999995	0.000004
83	PER004528	Paucarcolla	2	0.000459	0.999513	0.000029
84	PER004529	Paucarcolla	2	0.000547	0.993360	0.006093
85	PER004530	Paucarcolla	1	0.986019	0.000098	0.013883
86	PER004531	Paucarcolla	2	0.026495	0.967449	0.006056
87	PER004532	Yunguyo	3	0.053656	0.004471	0.941873
88	PER004533*	Yunguyo	3	0.083219	0.407678	0.509102
89	PER004534	Yunguyo	1	0.987744	0.001576	0.010679
90	PER004535*	Yunguyo	2	0.042889	0.751825	0.205286
91	PER004536*	Yunguyo	2	0.003988	0.668285	0.327726
92	PER004537	Yunguyo	3	0.000013	0.009619	0.990368
93	PER004538	Yunguyo	1	0.994631	0.000496	0.004873
94	PER004539	Yunguyo	3	0.013179	0.004102	0.982719
95	PER004540*	Paucarcolla	1	0.672758	0.141376	0.185866
96	PER004550	Pomata	3	0.000012	0.004810	0.995178
97	PER004551	Pomata	3	0.000125	0.005208	0.994667
98	PER004552	Juli	3	0.000003	0.000053	0.999944
99	PER004553	Juli	1	0.934612	0.002960	0.062428
100	PER004554	Pomata	3	0.000413	0.009049	0.990538
101	PER004555*	Pomata	3	0.275807	0.000093	0.724100
102	PER004556	Pomata	3	0.068472	0.000008	0.931520
103	PER004557	Pomata	3	0.002585	0.000082	0.997333
104	PER004558	Huacullani	3	0.000078	0.004816	0.995106
105	PER004559	Juli	3	0.000002	0.000019	0.999979
106	PER004560	Huacullani	3	0.000096	0.000007	0.999897
107	PER004561	Huacullani	3	0.000083	0.016655	0.983262
108	PER004562	Pomata	3	0.000577	0.000065	0.999357
109	PER004563	Pomata	3	0.004630	0.003147	0.992224
110	PER004564	Pisacoma	1	0.952283	0.041566	0.006151
111	PER004565	Pomata	3	0.028400	0.000137	0.971463
112	PER004566	Huacullani	2	0.000071	0.900545	0.099384
113	PER004567	Pisacoma	2	0.001133	0.996223	0.002644
114	PER004568	Pomata	2	0.001296	0.809729	0.188975
115	PER004569	Huacullani	3	0.093026	0.000053	0.906921
116	PER004570*	Pisacoma	3	0.008024	0.271573	0.720403
117	PER004571	Pomata	1	0.999686	0.000202	0.000113
118	PER004572	Pomata	3	0.030276	0.009556	0.960168
119	PER004573	Pomata	1	0.999672	0.000056	0.000271
120	PER004574	Pomata	2	0.000366	0.888290	0.111344
121	PER004575	Pomata	3	0.020771	0.010831	0.968398
122	PER004576	Pomata	1	0.919302	0.076833	0.003865
123	PER004578	Pomata	2	0.025241	0.927776	0.046983
124	PER004579	Paucarcolla	3	0.018458	0.059744	0.921798
125	PER004581*	Acora	3	0.037173	0.290610	0.672217
126	PER004582	Acora	1	0.862279	0.000283	0.137439

127	PER004583	Acora	2	0.000065	0.999886	0.000049
128	PER004584*	Pomata	3	0.138072	0.120477	0.741451
129	PER004585	Puno	2	0.000065	0.998755	0.001181

Tabla 6.4. Número total de alelos y heterocigosidad esperada de cada locus (Hexp), y número de alelos privados por grupo y locus. Los valores de Hexp fueron hallados utilizando la función *summary* del paquete *adegenet* en el software R. Los marcadores con un asterisco fueron altamente polimórficos de acuerdo a la definición de Ott (Ott, 1992).

Locus	Número total de alelos	Número de alelos por grupo						Hexp
		K1		K2		K3		
		Número de alelos privados	Número total de alelos	Número de alelos privados	Número total de alelos	Número de alelos privados	Número total de alelos	
QGA021	9	5	9	0	5	0	7	0.64
QGA024*	13	0	11	1	12	1	12	0.87
QAAT027*	10	0	7	2	7	3	9	0.74
QAAT112*	10	0	6	0	8	6	10	0.78
QAAT106*	10	0	6	1	8	3	7	0.73
QGA032	8	0	6	0	5	7	8	0.69
KAAT006	7	1	6	0	5	3	6	0.61
QAAT24	11	2	8	0	6	5	9	0.68
QAAT50*	16	0	10	1	12	3	14	0.82
KAAT037*	18	1	9	2	12	8	13	0.82
QAAT70*	9	0	7	2	7	0	7	0.77
KGA3*	10	0	5	1	9	2	9	0.76
QAAT76*	15	0	9	4	13	2	13	0.87
QAAT74*	14	1	11	0	10	4	13	0.86
QAAT100*	19	4	13	1	14	1	17	0.87
Total	179							
Promedio	11.93							0.77

Tabla 6.5. Número de accesiones, heterocigosidad esperada (Hs), número de alelos, número promedio de alelos por locus y número de alelos privados para los 3 grupos identificados por el DAPC.

Grupo genético	Número de accesiones	Hs	Número de alelos	Número promedio de alelos por locus	Número de alelos privados
K1	39	0.6957	123	8.2	14
K2	37	0.7198	133	8.87	15
K3	53	0.7761	154	10.27	48

Tabla 6.6. Estimaciones de los valores F_{ST} por pares (por debajo de la diagonal) y los valores G_{ST} por pares (por encima de la diagonal) entre los grupos hallados por el DAPC utilizando las frecuencias alélicas por el método de De Silva y colaboradores (De Silva et al., 2005) y la función *calcPopDiff* del programa Polysat.

Grupos	K1	K2	K3
K1	-	0.038462	0.044945
K2	0.041962	-	0.026379
K3	0.046165	0.028807	-

Tabla 7.1. Comparación del número de alelos (A) y heterocigosidad (H) de diferentes marcadores microsatélites de *Chenopodium quinoa* Willd. obtenidos de diferentes investigaciones. La lista de marcadores corresponde a los utilizados en el presente estudio, no se muestra valores de otros marcadores usados en las otras investigaciones que no se hayan utilizado en el presente estudio.

Locus	Mason et al. (2005)		Jarvis et al. (2008)		Christensen et al. (2007)		Fuentes et al. (2009)		Fuentes et al. (2012)		Costa et al. (2012) ^a		Costa (2014) ^{a,b}		Vía y Rada (2015)		Presente estudio	
	A	H	A	H	A	H	A	H	A	H	A	H	A	H	A	H	A	H
QGA021	9	0.85	-	-	-	-	-	-	-	-	-	-	-	-	9	0.58	9	0.64
QGA024	8	0.81	-	-	-	-	-	-	-	-	-	-	-	-	20	0.87	13	0.87
QAAT027	8	0.83	-	-	-	-	-	-	-	-	20	0.87	22	0.91	10	0.65	10	0.74
QAAT112	8	0.81	-	-	-	-	-	-	-	-	11	0.78	15	0.76	13	0.68	10	0.78
QAAT106	7	0.8	-	-	-	-	-	-	-	-	12	0.75	19	0.86	17	0.79	10	0.73
QGA032	6	0.81	-	-	-	-	-	-	-	-	-	-	-	-	11	0.59	8	0.69
KAAT006	-	-	6	0.79	-	-	-	-	-	-	-	-	-	-	9	0.34	7	0.61
QAAT24	6	0.58	-	-	20	0.85	14	0.83	8	0.84	19	0.84	25	0.92	13	0.8	11	0.68
QAAT50	9	0.87	-	-	27	0.89	10	0.82	8	0.74	16	0.84	31	0.92	19	0.83	16	0.82
KAAT037	-	-	13	0.9	-	-	-	-	-	-	-	-	-	-	31	0.85	18	0.82
QAAT70	7	0.77	-	-	17	0.91	11	0.79	7	0.74	-	-	-	-	12	0.84	9	0.77
KGA3	-	-	5	0.65	21	0.85	8	0.67	5	0.7	9	0.73	17	0.82	10	0.74	10	0.76
QAAT76	8	0.82	-	-	27	0.92	20	0.9	11	0.87	20	0.92	32	0.93	25	0.87	15	0.87
QAAT74	9	0.85	-	-	16	0.87	9	0.77	8	0.79	14	0.82	19	0.91	18	0.88	14	0.86
QAAT100	6	0.82	-	-	-	-	-	-	-	-	-	-	-	-	44	0.96	19	0.87

^a Se evaluó la variabilidad intra-accesión utilizando plantas individuales por accesión en vez de utilizar la técnica del bulk.

^b Los valores son para las accesiones solo de Sudamérica (Extra NOA) (Costa, 2014).

12.2. Figuras

Figura 2.1. Representación esquemática del evento de aloploidización en la evolución de la quinua (Bazile et al., 2014).

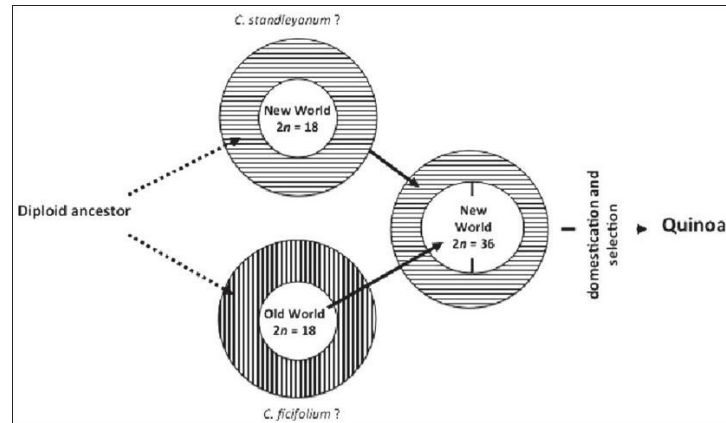


Figura 2.2. Distribución de la quinua desde tiempos precolombinos. El patrón más denso de puntos indica las zonas actuales de mayor producción (Extraído de Costa, 2014).



Figura 2.3. Muestras de quinuas de restos arqueológicos comparado con el cultivo actual (Extraído de Bazile et al., 2014)

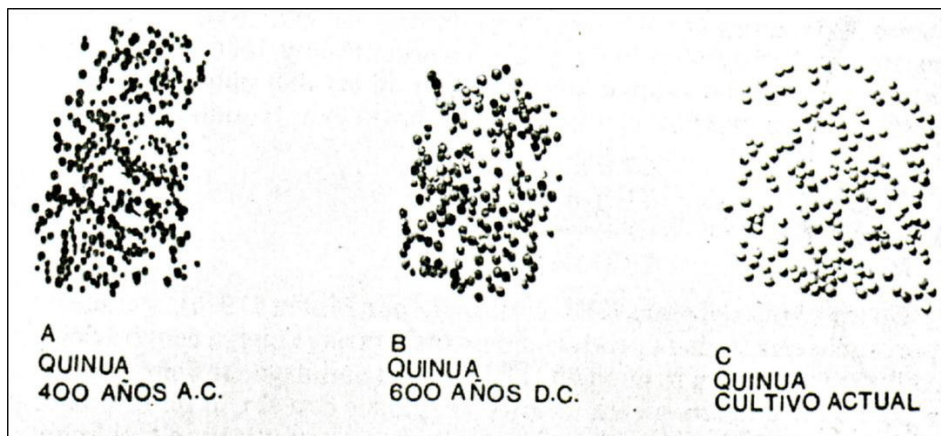


Figura 2.4. Hábitos de crecimiento de la quinua: a) Simple, b) ramificado hasta el tercio inferior, c) ramificado hasta el segundo tercio y d) ramificado con panoja principal no diferenciada (Extraído de Rojas et al., 2014).

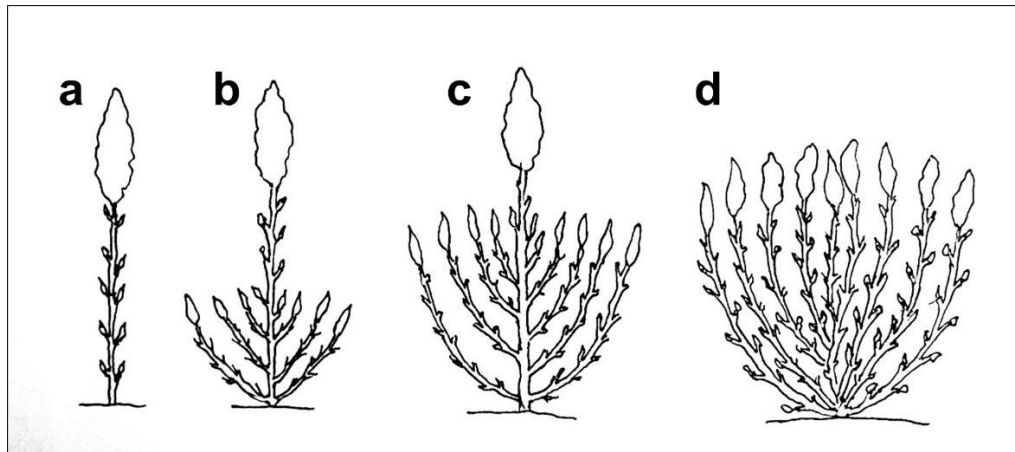


Figura 2.5. Forma de la panoja: a) Glomerulada, b) intermedia y c) amarantiforme (Extraído de Rojas et al., 2014).

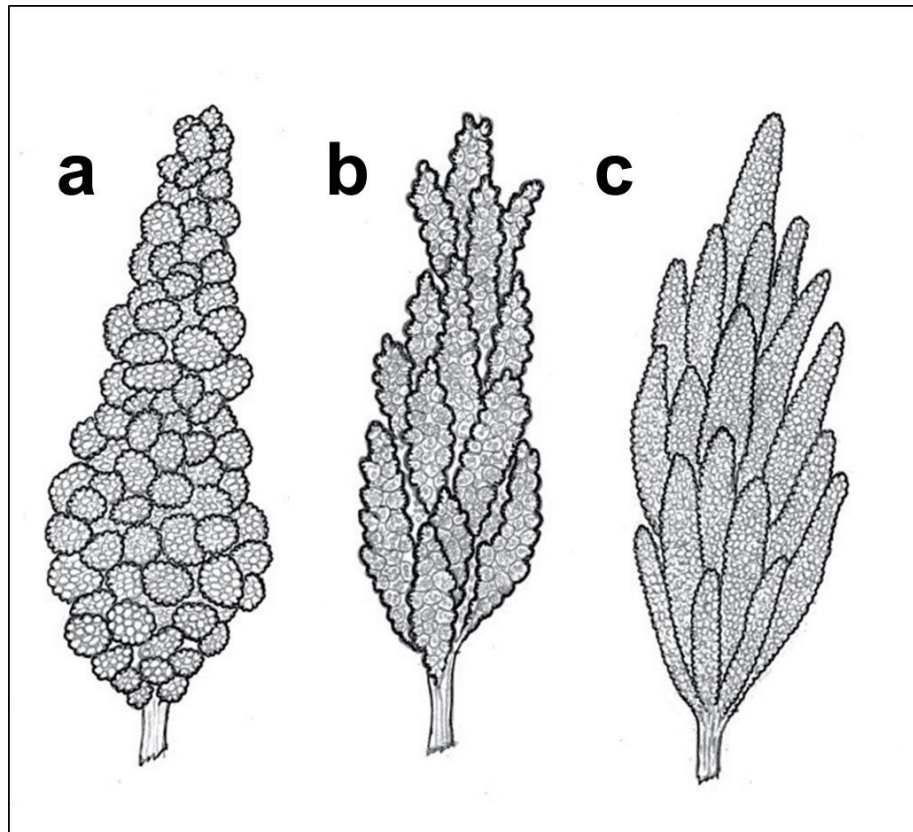


Figura 2.6. Formas de grano de quinua: a) lenticular, b) cilíndrica, c) elipsoidal y d) cónica (Extraído de Rojas et al., 2014).

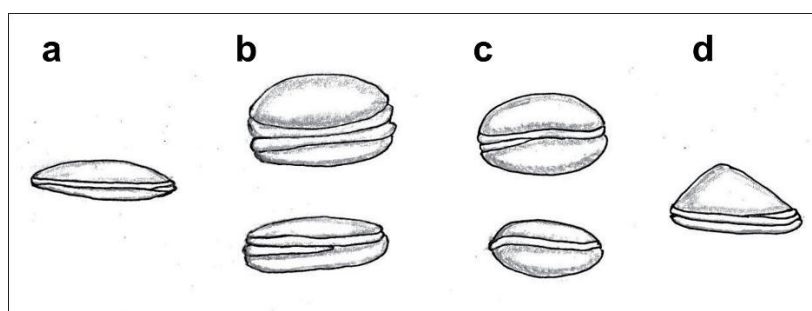


Figura 2.7. Distribución de los ecotipos de quinua en los sub-centros de diversidad: A. Valles Interandinos, B. altiplano, C. yungas, D. salares, y E. costa (Extraído de Bazile et al., 2014).

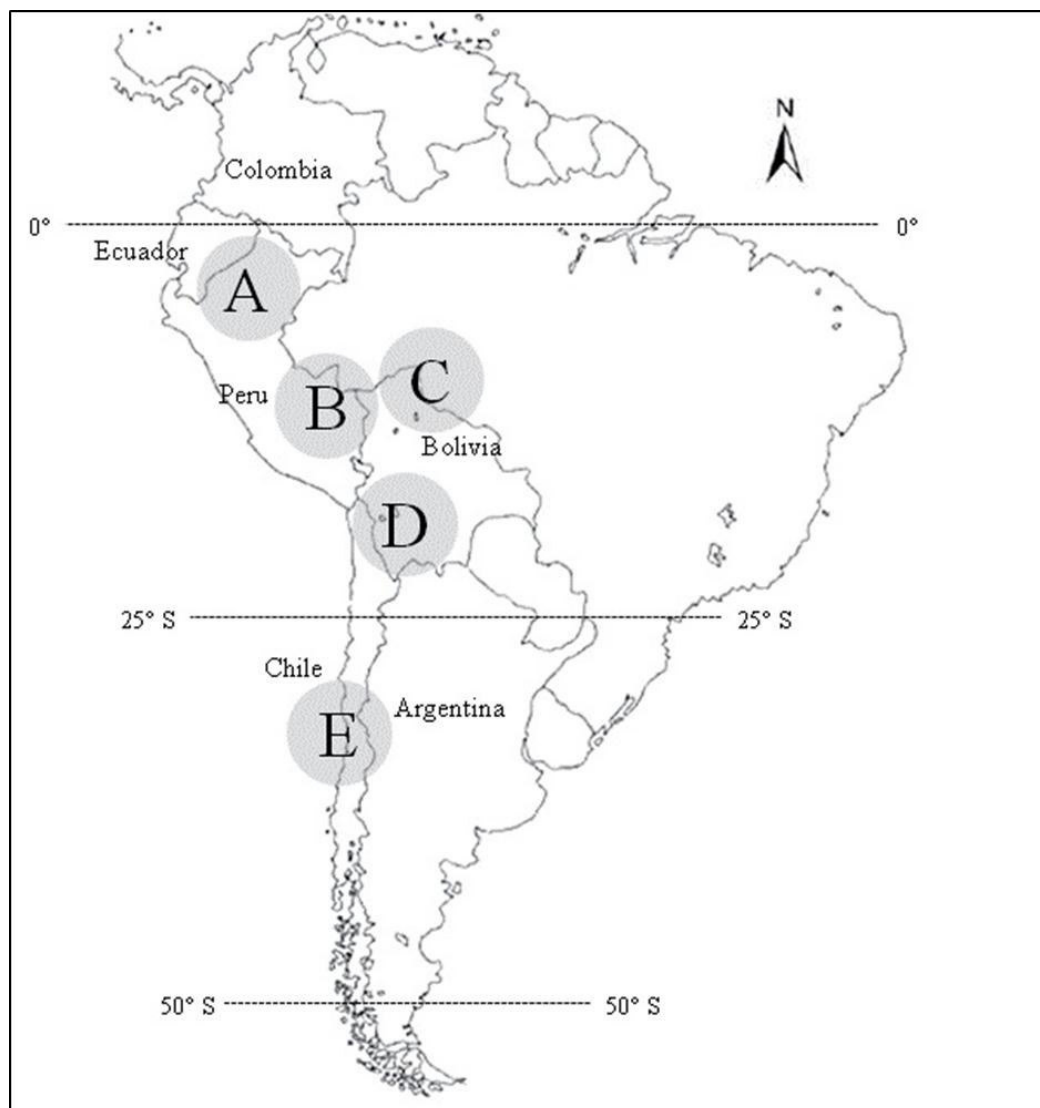


Figura 2.8. Esquema representando una región microsatélite como marcador molecular (Extraído de Costa, 2014).

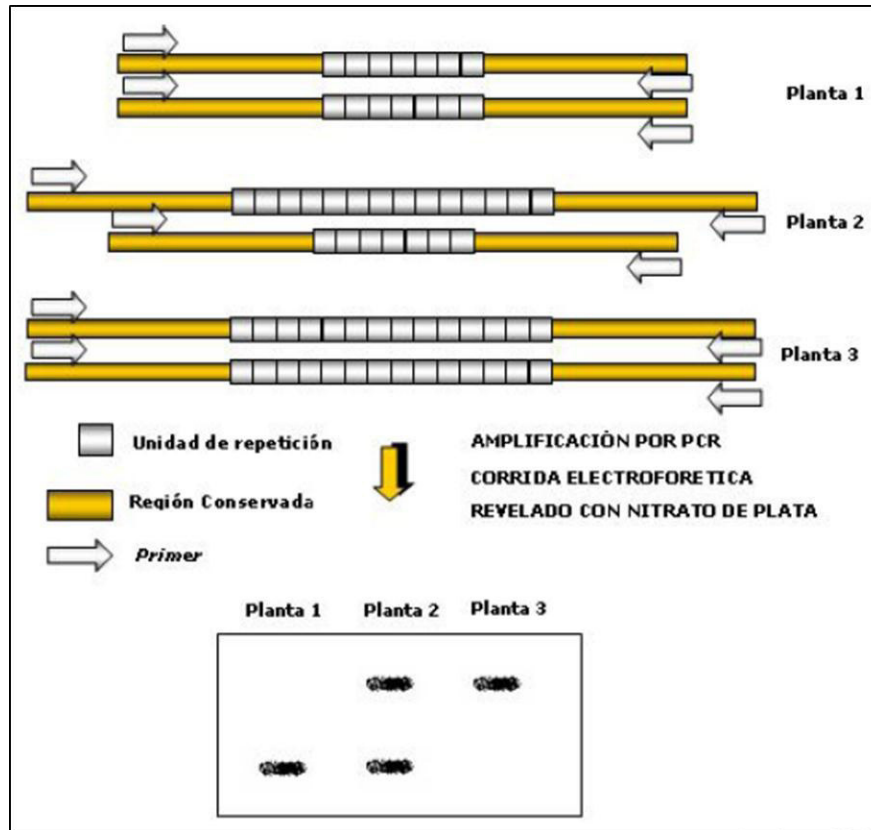


Figura 2.9. La replicación conlleva a la formación de nuevos alelos con menos o más repeticiones siempre y cuando una de las hebras contenga un error (Extraído de Madesis et al., 2013).

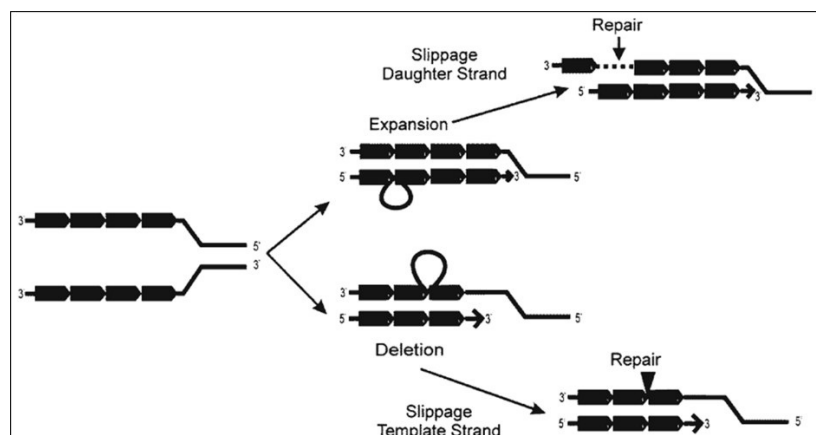


Figura 2.10. Esquema de los instrumentos de electroforesis capilar utilizados para el análisis de ADN (Extraído de Butler, 2005).

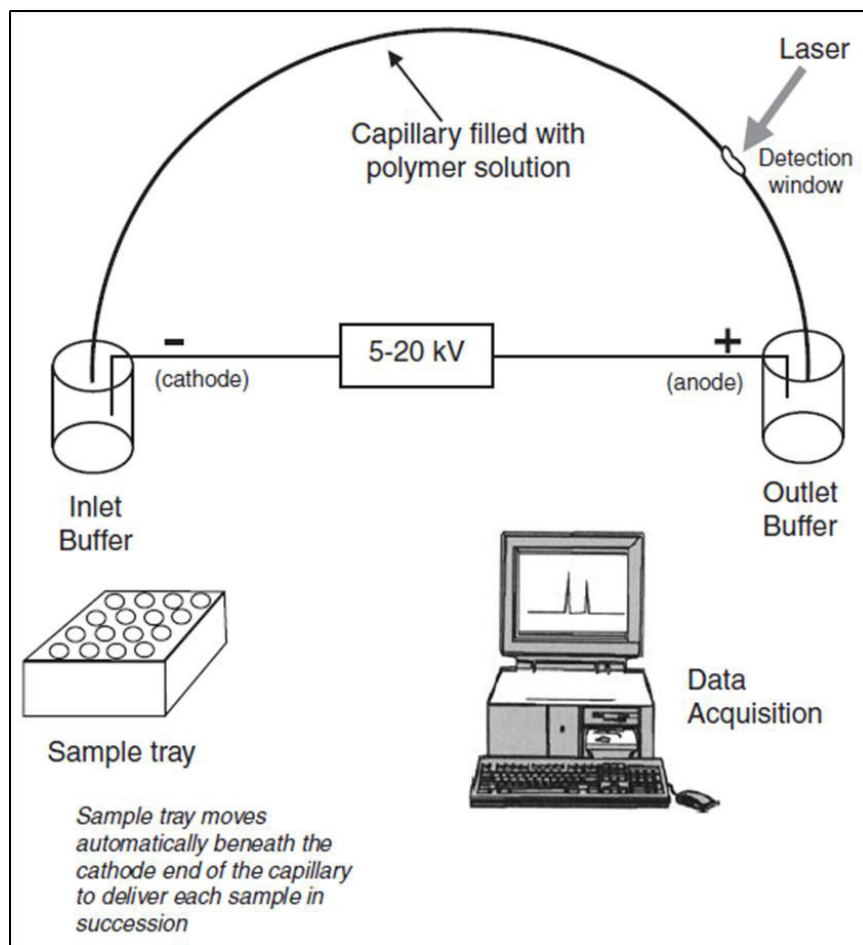


Figura 2.11. Diagrama del marcaje dual de cebadores. En el primer ciclo de PCR, una región microsatélite es amplificada usando un cebador directo que presenta una secuencia única en su extremo (*Forward tailed primer*) y un cebador reverso. La PCR continúa hasta el ciclo 30 cuando hay múltiples copias del mismo fragmento. En los ciclos 31-38, un cebador con la secuencia única unido a un fluoróforo es combinado con el cebador reverso para marcar fluorescentemente los mismos fragmentos. El ciclo 39 consiste de una extensión final para asegurar que la mayoría de fragmentos estén propiamente amplificados y marcados (Extraído de Culley et al., 2013)

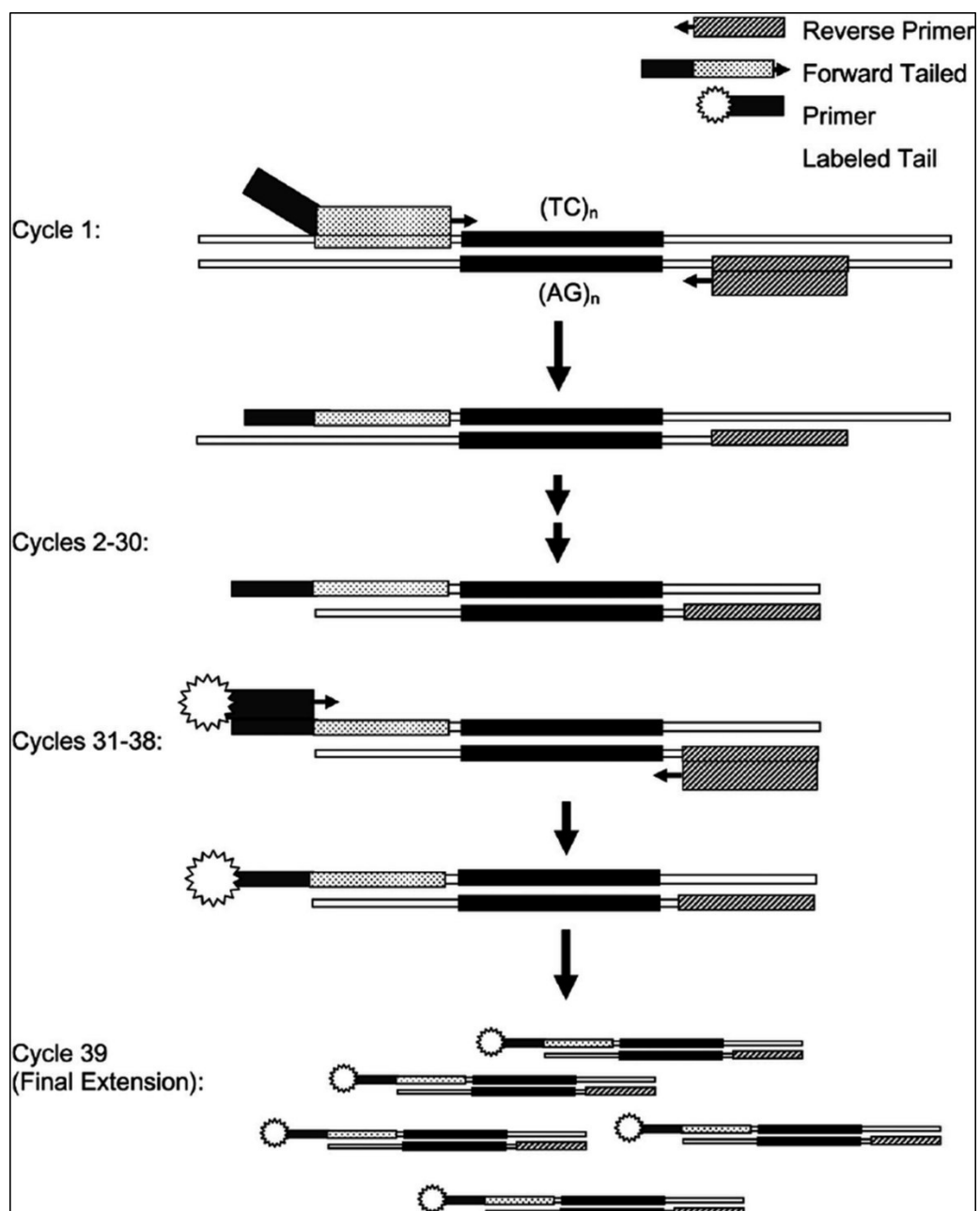


Figura 5.1. Ubicación Geográfica de las 22 localidades diferentes de las 129 accesiones de quinua, agrupadas en 13 distritos de 5 provincias (San Román, Puno, El Collao, Chucuito y Yunguyo) del sur de Puno. Mapa generado por Google Earth, 2017.



Figura 5.2. Laboratorio de Biología Molecular y Genómica en la sede central del INIA, distrito de La Molina, Lima, Perú.



Figura 5.3. Fotodocumentador Chemidoc™ XRS+ de Bio-rad (a) y analizador genético ABI 3130xl (b).

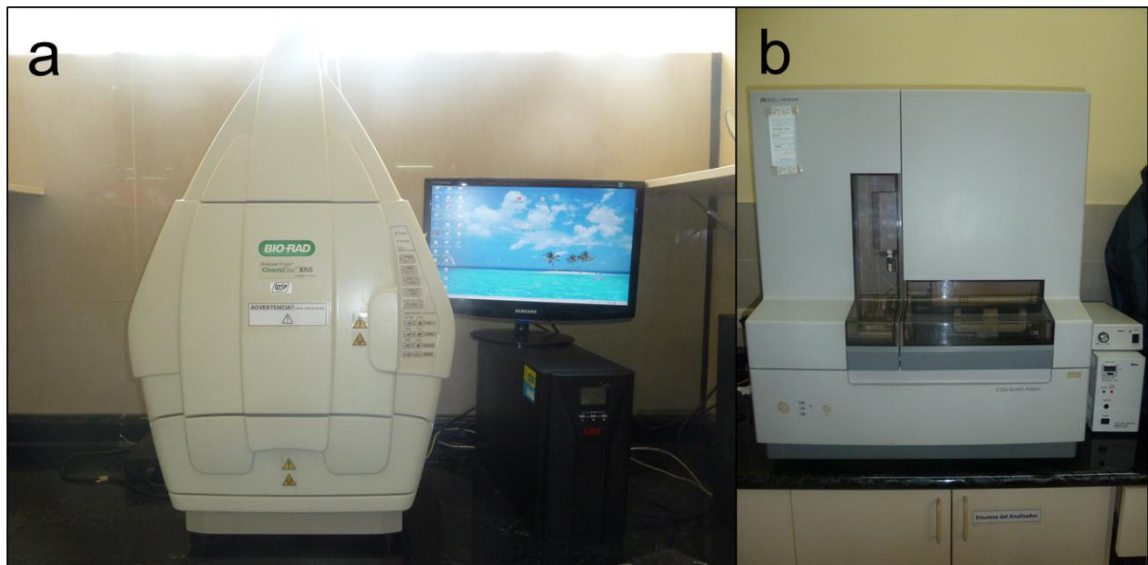


Figura 6.1. Gel de agarosa al 1% con ADN genómico de cada accesión de *Chenopodium quinoa* Willd. por pocillo.

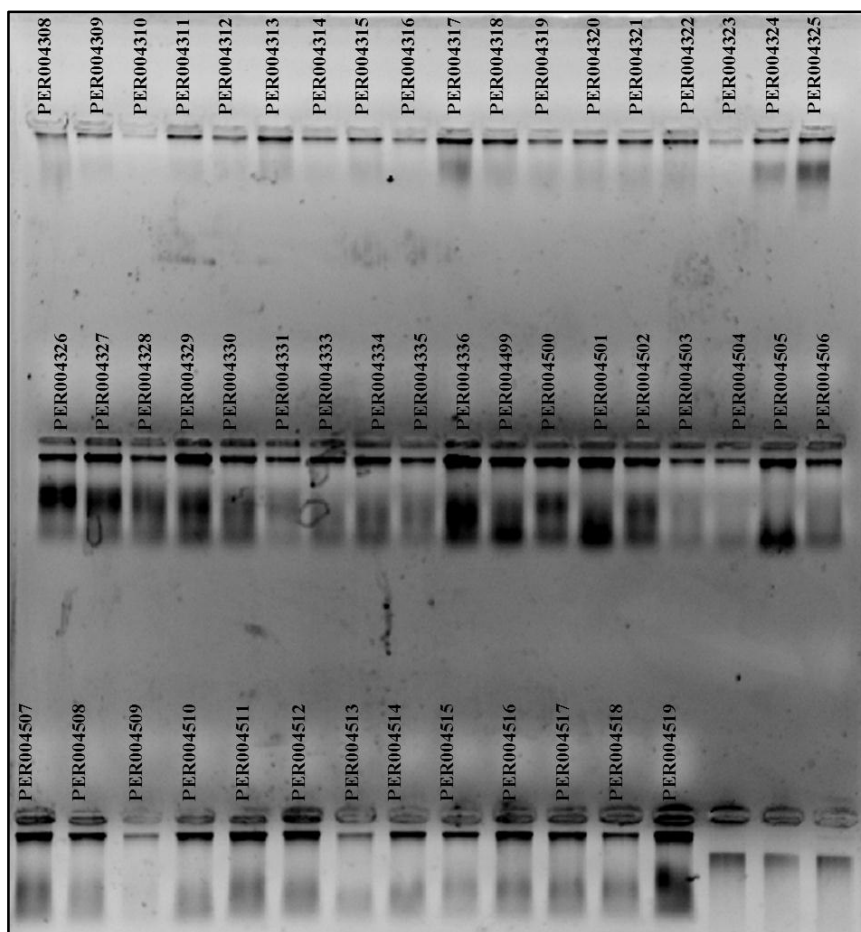


Figura 6.2. Gel de agarosa al 2% con regiones microsatélites amplificadas del genoma de *Chenopodium quinoa* Willd. Los marcadores microsatélites utilizados fueron QGA021/ QGA024 (1), QAAT027/QAAT112 (2), QAAT70/QAAT106 (3) y QGA032 (4). Las bandas del marcador de peso molecular (*DNA ladder*) se encuentran en pb.

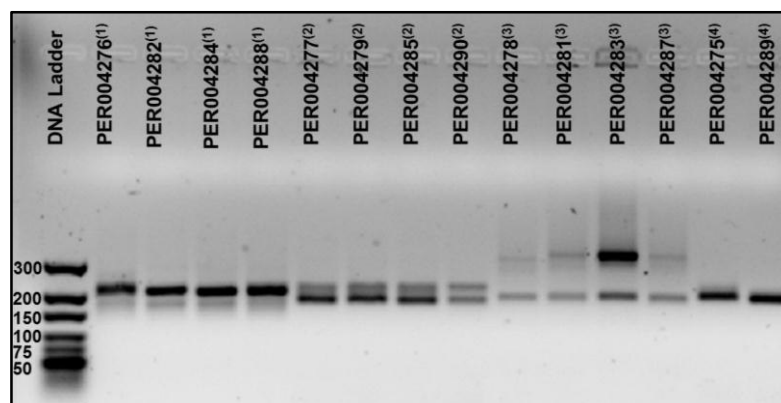


Figura 6.3. Electroferogramas generados por el software GeneMapper® versión 4.0. Se muestran los perfiles de los amplificados de los marcadores QGA021 (a), QAAT24 (b), QAAT106 (c) y QGA032 (d) para las accesiones PER004318 (a), PER004304 (b), PER004276 (c) y PER004328 (d) respectivamente. Las secuencias microsatélites dinucleótido (a y d) van acompañadas por más picos tartamudos a comparación de las secuencias microsatélites trinucleótido (b y c).

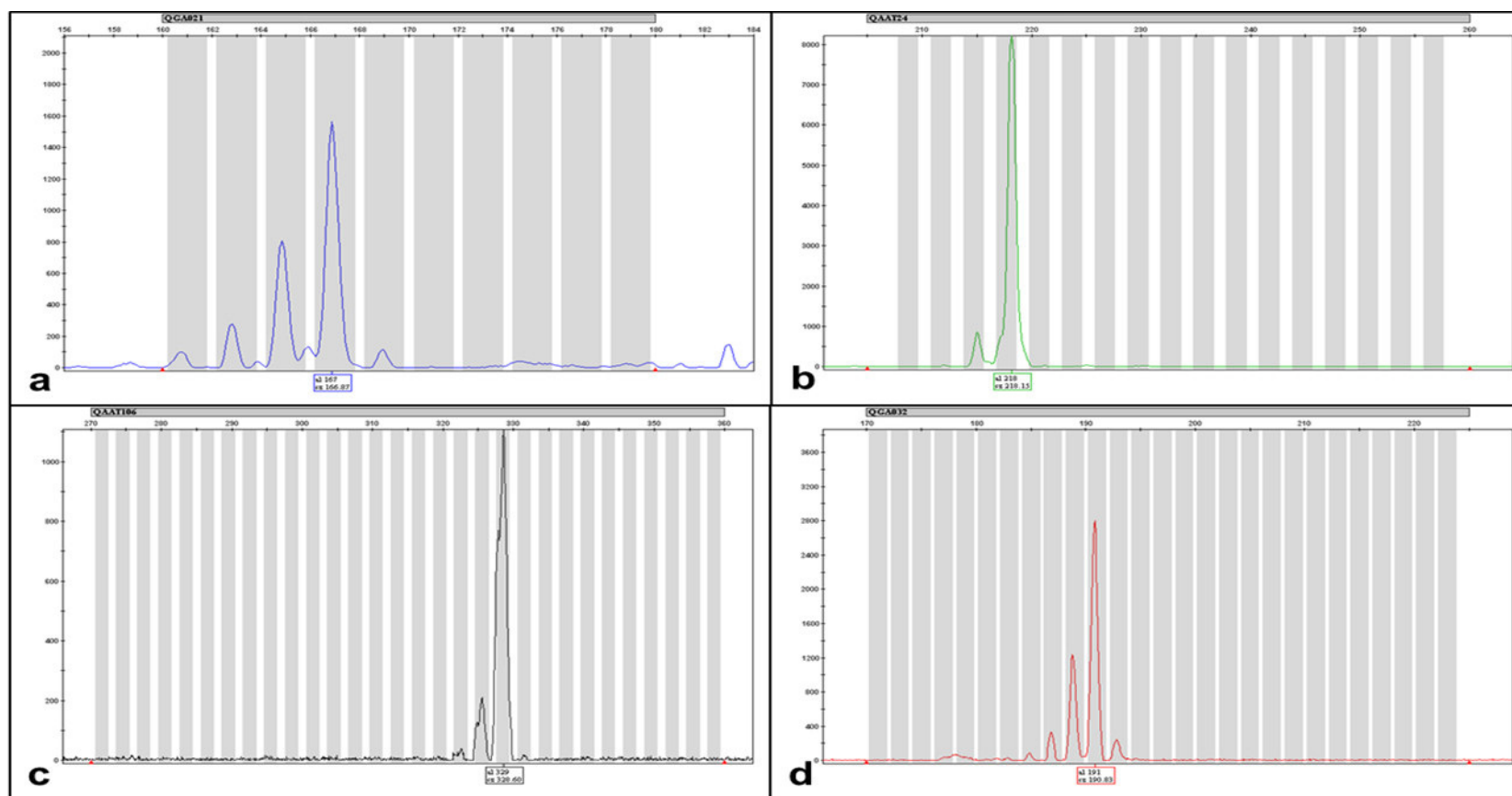


Figura 6.4. Resultados del análisis (segunda evaluación) de agrupamiento bayesiano utilizando STRUCTURE. Se evaluó valores de K del 1 al 7 (20 repeticiones por cada K), usando un modelo mixto, frecuencias alélicas correlacionadas, la función LOCPRIOR (información de las localidades), 10000 iteraciones burn-in y 50000 pasos de cadena de Markov Monte Carlo (MCMC). a) Variación de la probabilidad logarítmica de los datos en función al número de grupos (K). b) Variación de los valores ΔK utilizando el método de Evanno y colaboradores (Evanno et al., 2005). El número óptimo de grupos presenta el mayor valor de ΔK (K=3).

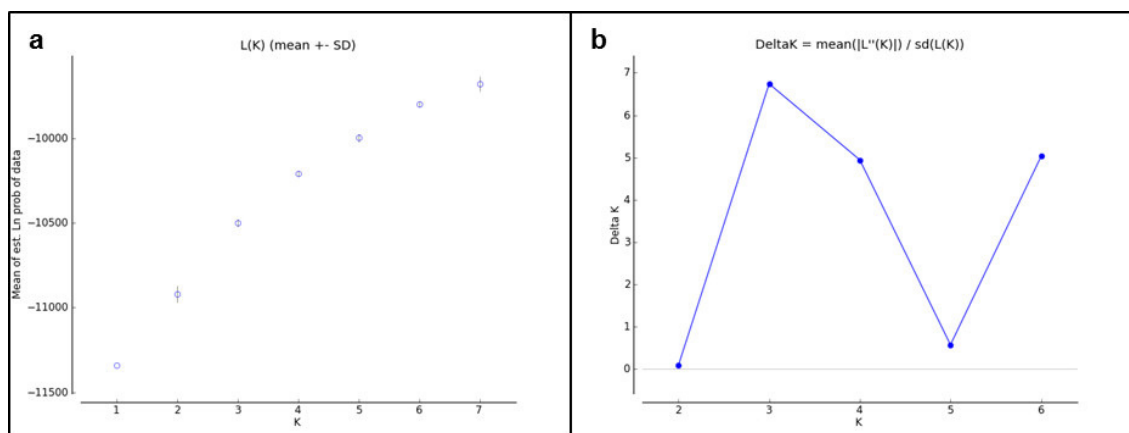


Figura 6.5. Resultados del agrupamiento por el método de k-medias utilizando la función *find.clusters*. Un valor de K de 3 (el menor valor del criterio de información bayesiano, BIC) representa el agrupamiento más óptimo de los datos.

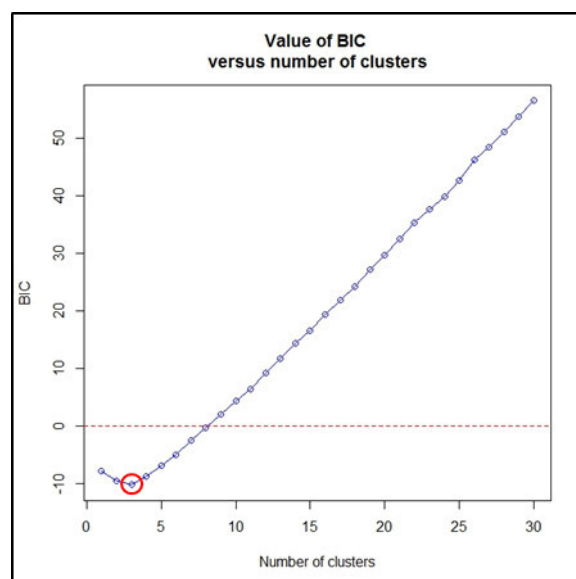


Figura 6.6. Tamaño de los grupos hallados por las funciones *find.clusters* (a) y *dapc* (b) según las probabilidades previas y posteriores respectivamente, y porcentaje de reasignación de las accesiones a cada grupo utilizando la función *dapc* (c).

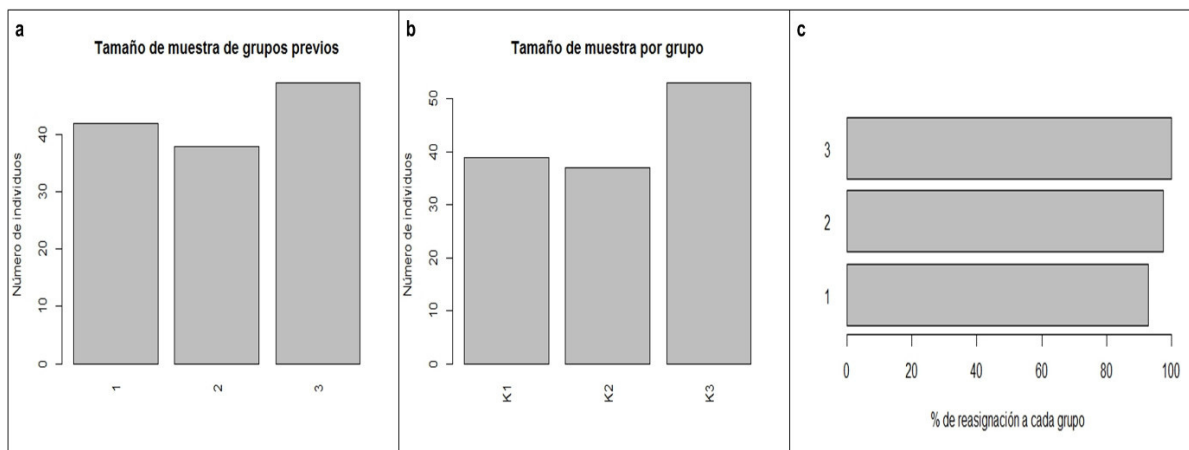


Figura 6.7. Gráfico de los resultados de la optimización del valor α -score utilizando la función *optim.a.score*. El valor óptimo del α -score se obtuvo reteniendo un número óptimo de 7 componentes principales.

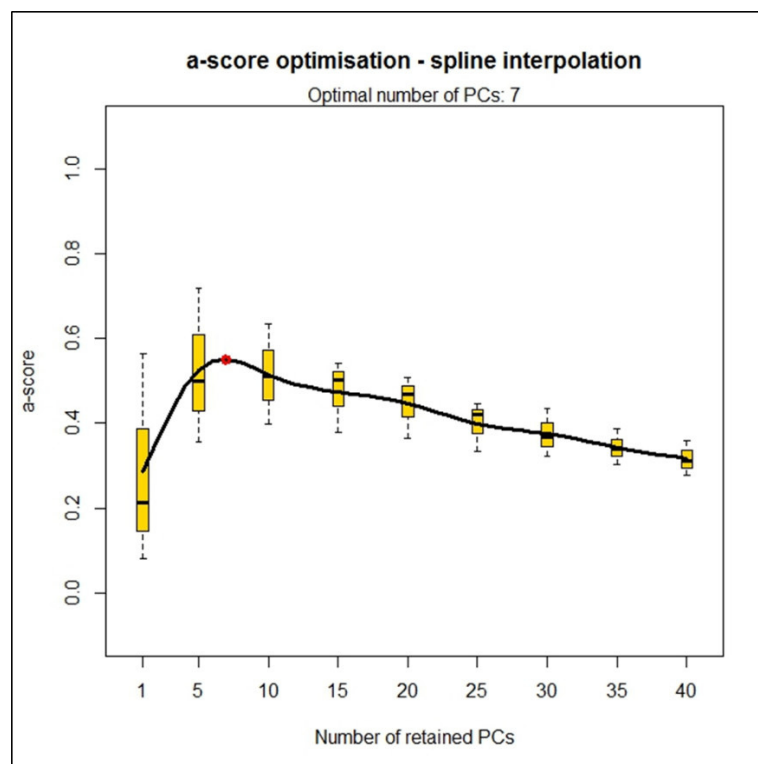


Figura 6.8. Gráfico de las 129 accesiones de quinua en los dos primeros factores discriminantes del análisis discriminante de componentes principales (DAPC) utilizando los 3 grupos hallados con la función *find.clusters* (K1, K2 y K3). El gráfico representa a las accesiones como cuadrados, triángulos y círculos encerrados en una elipse con un 95% de confianza. Los eigenvalores del análisis discriminante y del análisis de componentes principales representan la cantidad de variación genética capturada por el análisis del DAPC.

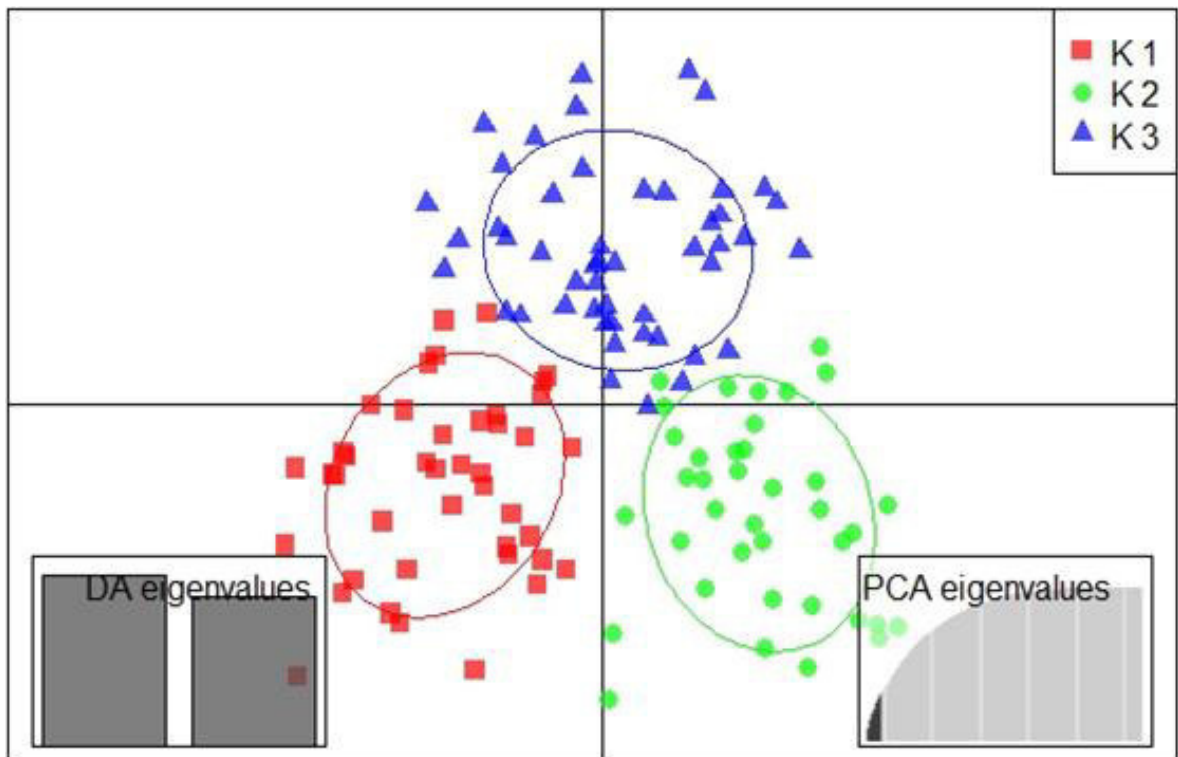


Figura 6.9. Frecuencias relativas de la procedencia (por distrito) de las accesiones no mixtas de *Chenopodium quinoa* Willd. en cada grupo determinado por el DAPC. Los tres grupos genéticos fueron los grupos K1 (a), K2 (b) y K3 (c).

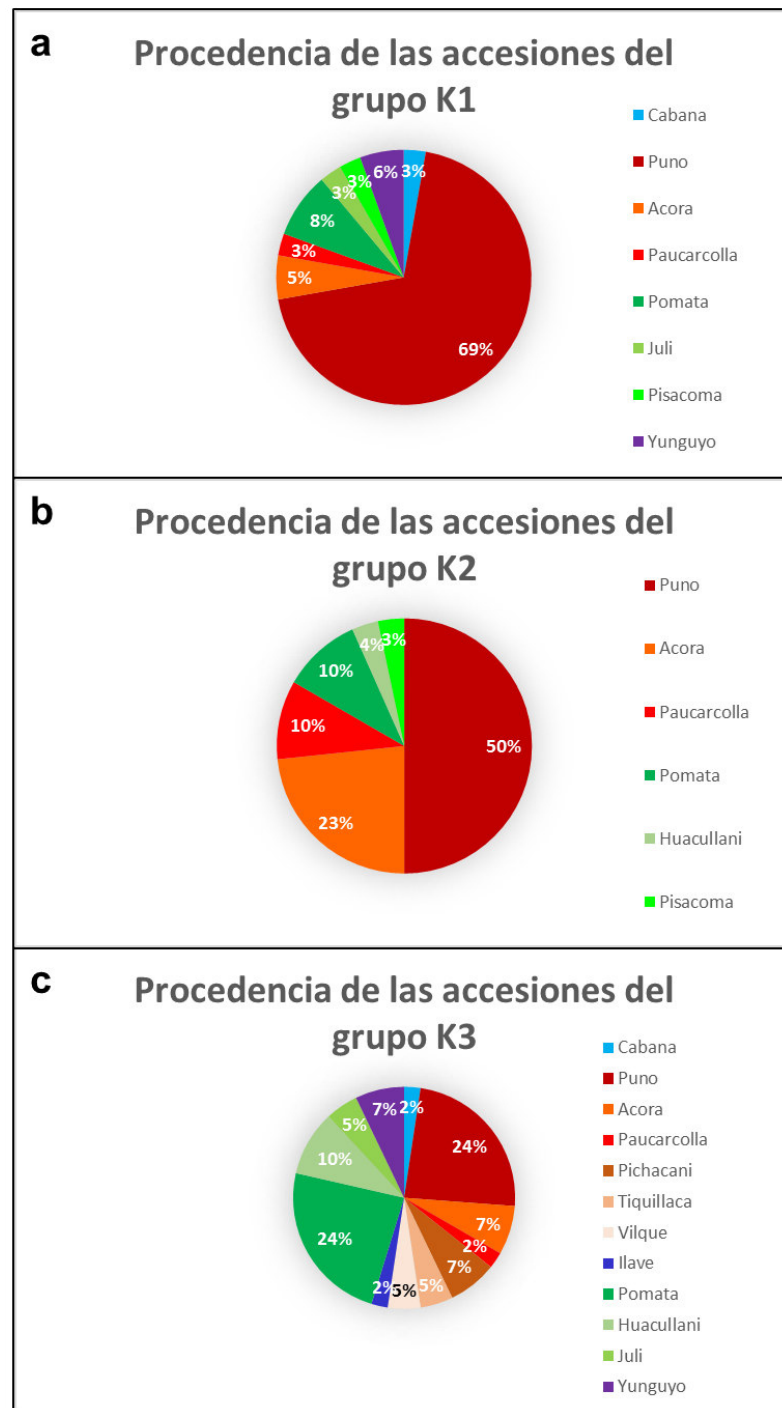
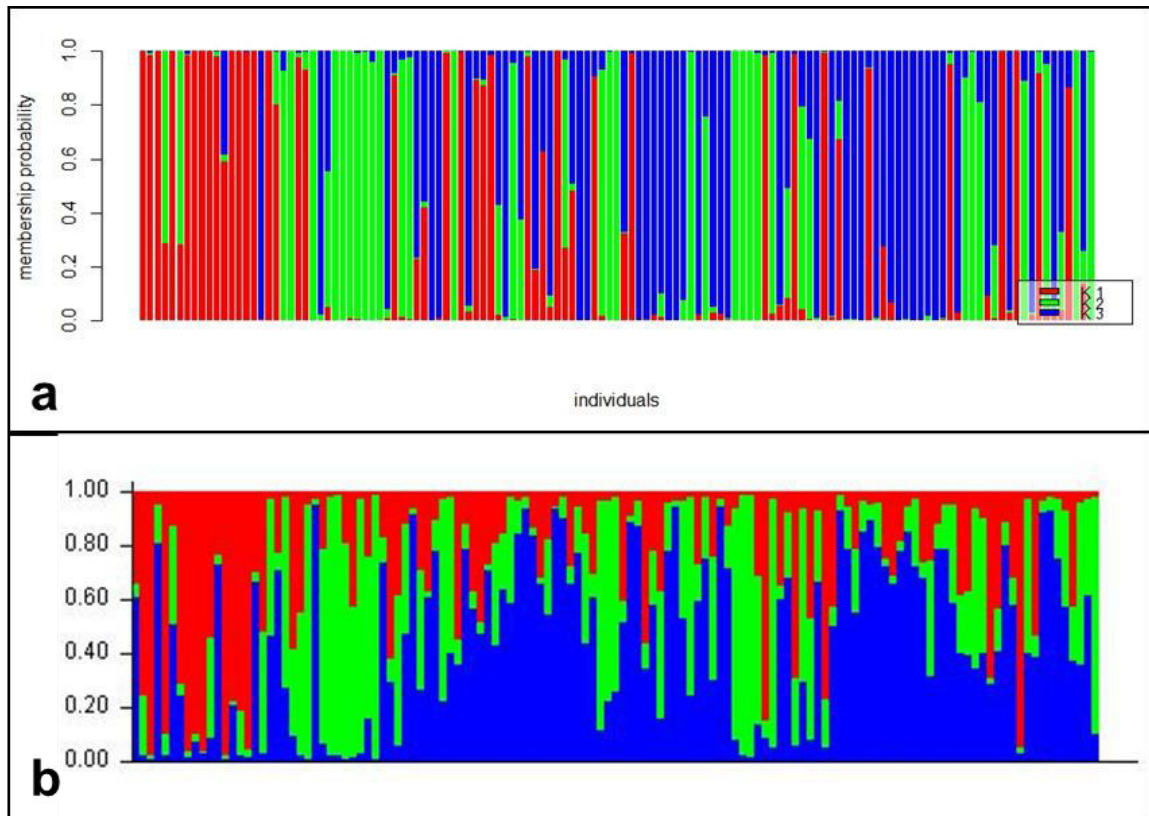


Figura 6.10. Gráficos de las probabilidades de membresía de las 129 accesiones de quinua en 3 grupos (K=3) utilizando a) DAPC y b) STRUCTURE. Cada individuo es representado por una barra vertical, con los colores correspondiendo a las probabilidades de membresía a los grupos K1 (rojo), K2 (verde) y K3 (azul).



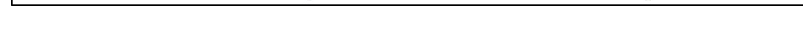


Figura 6.12. Valores P para las correlaciones negativas entre los alelos del locus QAAT100 del grupo K1. La función *alleleCorrelations* del paquete Polysat en el software R fue utilizada para la obtención de este gráfico. Los valores P varían del 0 (rojo) al 1 (blanco). Se asume herencia polisómica del locus debido a que los valores P no dividen a los alelos en dos grupos (isoloci).

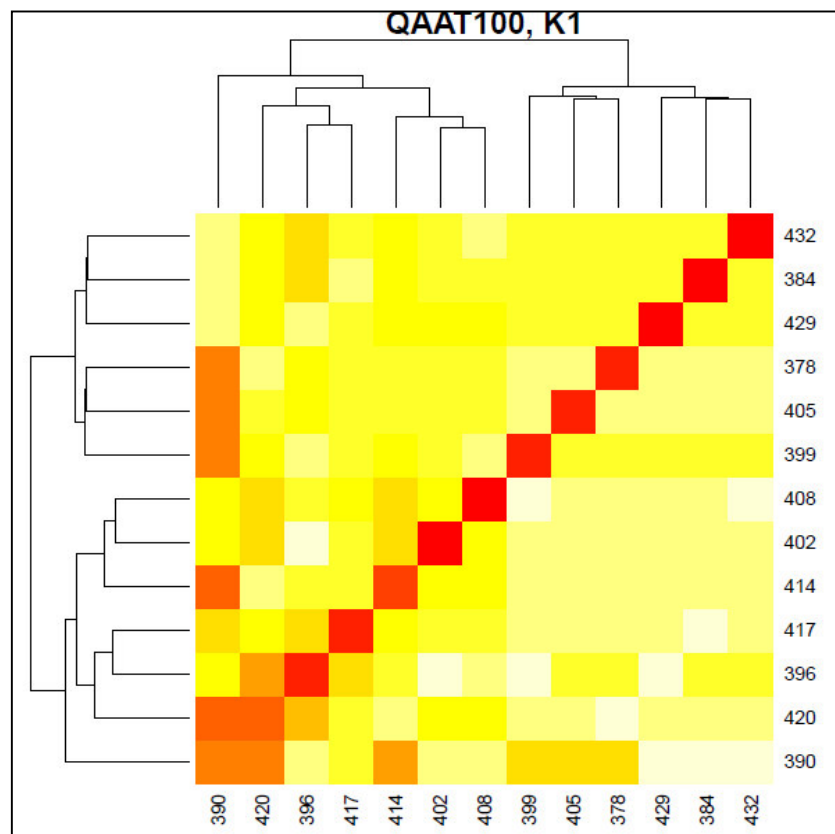


Figura 6.13. Gráfico del análisis de coordenadas principales (PCoA) de las 129 accesiones de quinua basado en la distancia de Bruvo (Bruvo et al., 2004) y las frecuencias alélicas por el método de De Silva y colaboradores (De Silva et al., 2005). Los colores rojo, verde y azul representan la pertenencia de cada accesión a cada grupo hallado en el DAPC (K1, K2 y K3), respectivamente.

